

PRO GRADU -TUTKIELMA

Taru Vihinen

Vakuutusyhtiön korkoriskin arviointi dynaamisen
Nelson-Siegel-mallin avulla

TAMPEREEN YLIOPISTO

Informaatiotieteiden yksikkö

Tilastotiede

Joulukuu 2014

Tampereen yliopisto

Informaatiotieteiden yksikkö

VIHINEN, TARU: Vakuutusyhtiön korkoriskin arviointi dynaamisen Nelson-Siegel-mallin avulla

Pro gradu -tutkielma, 57s., 2 liites.

Tilastotiede

Joulukuu 2014

TIIVISTELMÄ

Vakuutusyhtiöt hyödyntävät kuolevuusennusteita muun muassa tuotteidensa hinnoitteluun ja tarkasteluun. Kuolevuudella tarkoitetaan kuolleiden suhteellista osuutta tietystä ikäluokasta. Kuolevuudella on jo useiden vuosikymmenten ajan havaittu olevan laskeva trendi, toisin sanoen elinajanodotteet kasvavat vuosi vuodelta niin naisten kuin miestenkin keskuudessa. Parhaan kuolevuusennusteen saa käyttämällä kohorttikohtaista kuolevuustaulua, mutta usein kuolevuus kuitenkin esitetään funktiomuodossa. Tässä tutkielmassa sovite muodostetaan naisten kohorttikohtaisiin henkivakuutuskuolevuuksiin ryhmäeläkevakuutuksen osalta käyttämällä Makehamin kuolevuusmallia. Funktio muodostetaan käyttäen sekä lineaarista että epälineaarista regressiota.

Korkoriskit ovat usein merkittäviä henkivakuutusyhtiöille. Jotta päästään mallintamaan korkoriskiä, muodostaan tutkielmassa kuolevuusmallin lisäksi dynaaminen Nelson-Siegel-korkomalli Yhdysvaltojen liikkeelle laskemien joukkovelkakirjojen markkinatuottojen koroille käyttäen AR(2)-prosessia. Korkomalli koostuu kolmesta estimoidusta faktorista sekä vakioidusta aikafaktorista. Estimoiduille faktoreille sekä näistä faktoreista muodostuvalle tuottokäyrälle muodostetaan lopuksi ennusteet simuloimalla.

Eläkettä voi saada korkeintaan niin kauan kuin siihen oikeutettu henkilö on elossa, minkä vuoksi käytettävän koron lisäksi eläkkeen pääoma-arvo riippuu myös kuolevuudesta. Tutkielmassa lasketaan kuolevuus- ja korkomallin avulla elinikäisen eläkkeen pääoma-arvoja valitulle kohortille. Lisäksi tutkielmassa tarkastellaan vakavaraisuuspääomavaatimusta (Solvency Capital Requirement) niin kutsutusta varojen ja vastuiden hallinta (Asset Liability Management) -näkökulmasta tutkimalla Value at Risk (VaR) -luvun muuttumista eripituisilla korkosijoituksilla.

Asiasanat Makeham-kuolevuusmalli, regressio, Nelson-Siegel-korkomalli, AR(2)-prosessi, elinikäisen eläkkeen pääoma-arvo

SISÄLTÖ

1	Johdanto.....	3
2	Aineisto.....	5
2.1	Kuolevuusaineisto.....	5
2.1.1	Populaation henkivakuutuskuolevuus.....	7
2.1.2	Lukumäärä- ja summakuolevuus	8
2.1.3	Korjauskerroin	10
2.1.4	Ryhmäeläkevakuutus	12
2.2	Korkoaineisto	13
3	Tutkimuksessa käytetyt menetelmät.....	15
3.1	Regressioanalyysi.....	15
3.1.1	Lineaarinen regressio	15
3.1.2	Epälineaarinen regressio	16
3.2	Aikasarja-analyysi.....	17
3.2.1	Stationaarinen prosessi.....	17
3.2.2	Otosauto- ja otos-osittaisautokorrelaatiofunktio.....	20
3.2.3	Jäännössarjan valkoisuuden testaaminen	22
4	Tutkimuksessa käytetyt mallit.....	25
4.1	Kuolevuusmallit.....	25
4.2	Korkomallit	27
4.2.1	Dynaaminen Nelson-Siegel-korkomalli.....	28
4.2.2	Nelson-Siegel-korkomalli	32
4.3	Solvenssi II.....	33
4.3.1	SCR -pääomavaatimus.....	33
4.3.2	Value at Risk.....	34
4.4	Elinikäinen eläke.....	35
5	Aineistojen analysointi.....	39

5.1	Makeham-kuolevuusmalli.....	39
5.2	DNS-mallin estimointi ja ennustaminen	41
5.3	Elinikäisen eläkkeen pääoma-arvot	53
5.4	VaR-luku.....	54
6	Johtopäätökset.....	56
	Lähteet.....	57
	Liitteet	60

1 JOHDANTO

Yleisesti kuolevuudella tarkoitetaan elossa olevan henkilön iästä riippuvaa todennäköisyyttä kuolla eli niin kutsuttua *kuolevuusintensiteettiä*. Aina varhaislapsuuden jälkeen kuolevuus on iän suhteen kasvava vaikkakin viime vuosikymmenten aikana kuolleisuuden suhteellinen pieneneminen on ollut nopeinta imeväisiässä. Lisäksi myös myöhemmän lapsuusiän kuolleisuus on vähentynyt. (Martelin, Koskinen & Valkonen, 2005.) Kuolevuudessa voi iän lisäksi olla eroja väestön eri populaatioiden välillä. Aineistojen lisäksi luvussa 2 esitellään elinajanodotteita ja niiden kehitystä. Syntymävuoteen pohjautuvan kuolevuuden välisen erottelun lisäksi jaottelu voitaisiin suorittaa sukupuolittain tai eri kehitystasolla olevien maiden populaatioiden välillä, sillä myös maiden välillä kuolevuuksissa on huomattavia eroja. Tässä työssä jaottelu tehdään kuolevuustrendin huomioonottamiseksi syntymävuoden perusteella. Lisäksi tarkastelun kohteena on vain aineiston naispopulaatio, sillä etenkin miehillä varallisuus vaikuttaa eläkevakuutuskuolevuuteen ja näin ollen miehet on päätetty jättää tarkastelun ulkopuolelle. Työn ensimmäisenä tavoitteena on tehdä funktiosovite K2012-raportin kuolevuuksille Lee-Miller-ennusteeseen naisten ryhmäeläkevakuutuksen tapauksessa.

Yksittäisten syiden vaikutusta kuolevuuteen ei pyritä tarkastelemaan, vaan sovitetta tehtäessä keskitytään yksinkertaisiin menetelmiin. Näillä menetelmillä on käytännössä päästy keskimäärin samoihin tuloksiin kuin ennusteilla, joissa on otettu huomioon erilaisia kuolinsyitä. Valituissa menetelmissä ei myöskään pyritä rajoittamaan kuolevuuden kehitystä kiinnittämällä elinajalle yläraja. Tähän on päädytty, sillä maailman väestön ikääntymisessä ei viimeiseen vuosisataan ole ollut viitteitä siitä, että elinajan kasvamiselle olisi lyhyellä aikavälillä joitakin rajoitteita. (Mäkinen, 2004.)

Kolmannessa luvussa käydään läpi tutkielmassa käytettäviä menetelmiä, eli regressio- ja aikasarja-analyysiä. Näistä ensimmäistä hyödynnetään sekä kuolevuus- että korkomallin muodostamisessa. Aikasarja-analyysiä hyödynnetään enimmäkseen korkomallissa mutta myös kuolevuusmallin ekstrapoloinnissa. Neljäs luku sisältää tutkielmassa käytettävien mallien esittelyn ja syvemmän tarkastelun.

Henkivakuutusyhtiöille korkoriskit ovat usein merkittäviä. Riskejä mitataan esimerkiksi ennakoivan vakavaraisuuden tai Solvenssi II:n mukaisissa vakavaraisuuskehikoissa. Jos yhtiö kuitenkin kaipaa tarkempaa käsitystä korkoriskeistä, on yhtiön perusteltua analysoida koron vaihtelua sekä sen mukanaan tuomia riskejä itse. Tällöin yhtiö voi hyödyntää esimerkiksi stokastisten aikarakenneyhtälöiden avulla saatuja korkomalleja. (Saraste, 2009.) Tutkielman toinen tavoite on korkomallin muodostaminen valit-

tuun korkoaineistoon sekä sen ennustaminen tulevaisuuteen simuloimalla. Lopullinen korkomalli muodostetaan luvussa 5.

Eläkevakuutuksen keskeisimmät riskit johtuvat sijoitusten tuottojen sekä kuolevuuden kehityksestä. Siis kun ollaan tekemisissä eläkevakuutuksen riskienhallinnan kanssa, ei riitä tarkastella vain pitkäikäisyyttä. Toinen tärkeä elementti koskee pitkäikäisyysriskin rahoittamista. Vakuutusyritysten on arvioitava velkansa ja niillä on oltava riittävästi varallisuutta sen kattamiseen. Lisäksi vakuutusyrityksillä on oltava pääomaa riskien kohtaamiseen. Simulointimallin avulla voidaan generoida stokastisia eläkkeiden kassavirtoja sekä analysoida pitkäikäisyysriskin rahoitusta eläkevakuutuksessa. (Ronkainen, 2012.) Tutkielmassa tätä tarkastellaan luvussa 5 käsitellen yhtä annuiteettikohorttia. Simulointituloksien lisäksi lasketaan vielä VaR-lukuja vakuutusyhtiön oman varallisuuden (Own Funds, *OF*) muutokselle vuoden aikana eripituisille korkosijoituksille.

Tutkielma jakaantuu neljään osaan, joista ensimmäisessä kuvaillaan aineistoja ja kuolevuusmallissa käytettävää ryhmäeläkevakuutusta, toisessa esitellään tutkielmassa käytettävät menetelmät, kolmannessa taas tutkielmassa käytettävät mallit. Neljännessä osassa lopulta muodostetaan sekä kuolevuus- että korkomalli, ja näiden mallien perusteella lasketaan elinikäisen eläkkeen pääoma-arvoja sekä tarkastellaan lisäksi vakavaraisuus- eli SCR-pääomavaatimusta eripituisilla korkosijoituksilla. Tutkielman aineistojen analysointi on pääosin toteutettu R-ohjelmistolla (R Core Team, 2013).

2 AINEISTO

Yhtenä tutkielman aineistona käytetään raportin K2012 (Sirén, 2012) referenssikuolevuudessa esitettyjä kuolevuuksia henkivakuutetuille vuosilta 1982–2012. Raportin kuolevuudet perustuvat tasoitettuun HMD-aineistoon (Human Mortality Database). Aineistossa vanhimpien ikäluokkien kuolevuuksia on tasoitettu, sillä näissä kuolevuuksissa vuosittainen vaihtelu on suurta populaation pienen koon vuoksi (Sirén, 2012). Jotta kuolevuustrendi voitaisiin ottaa parhaiten huomioon, on väestö sukupuolijaon lisäksi jaettu osapopulaatioiksi syntymävuoden perusteella (Kauppi, 2004).

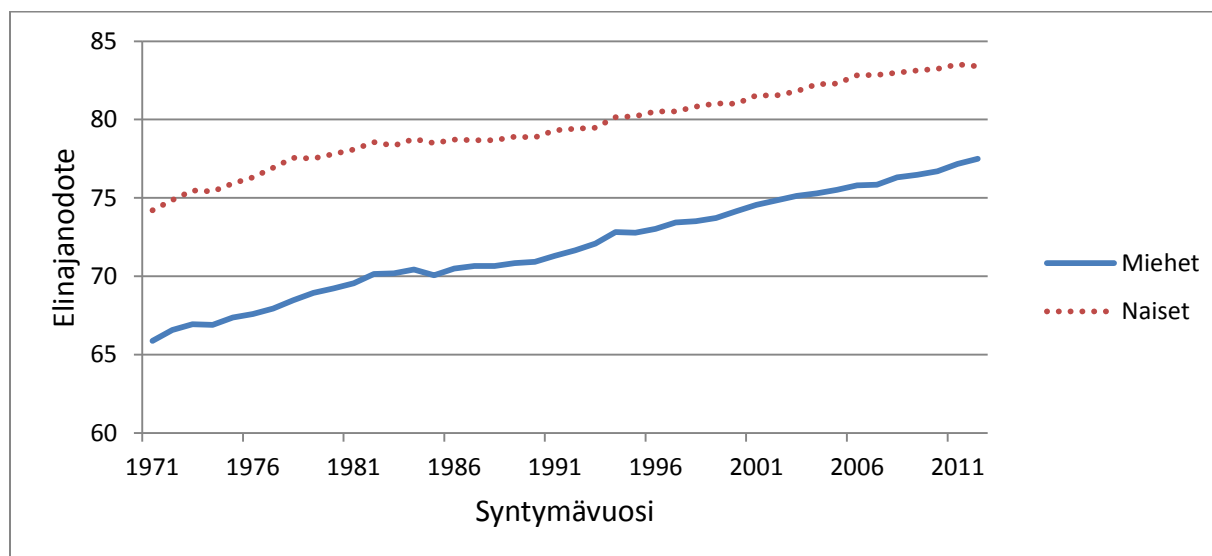
Toisena tutkielman aineistona toimii korkoaineisto The Federal Reserve -sivustolta. Korkoaineisto mittaa investointien perusteella noteerattuja valtion joukkovelkakirjojen tuottoja Yhdysvalloissa kiinteillä maturiteeteilla.

2.1 Kuolevuusaineisto

Tutkielmassa käytettävä kuolevuusaineisto on jaettu kahteen osaan: periodi- ja kohortikohtaiseen aineistoon. Jälkimmäisellä tarkoitetaan populaatiota, jonka jäseniä yhdistää sama syntymävuosi eli toisin sanoen sen jäsenet kuuluvat samaan syntymävuosikohorttiin. Kohortin seuraamiseen tarvitaan siihen saakka aikaa, että viimeinenkin samana vuonna syntynyt henkilö on kuollut. Mitä tuoreempi kohortti on kyseessä, sitä enemmän tehtävä kuolevuusmalli täytyy perustaa aikaisempien syntymävuosikohorttien kehittymisen perusteella tehtäviin ennusteisiin. Periodin eli tietyn kalenterivuoden populaatio taas on poikkileikkaus eri syntymävuosikohorteista. Sen tarkoitus on kuvata väestöä yhden kalenterivuoden aikana syntymävuodesta riippumatta. (Kauppi, 2004.) Periodikuolevuus määritellään aina periodin populaatiolle, eikä se siis toisin sanoen sisällä tietyn henkilön kuolevuuden kehitystä. Tässä tutkielmassa käytetään kuitenkin vain kohortikohtaista aineistoa. Funktiosovitetta tehtäessä sekä mallien sopivuutta tarkasteltaessa aineistona toimii kohorttiaineisto vuosilta 1960–2012 sekä kertoimille valittujen mallien avulla ekstrapoloitu aineisto vuodesta 1940 vuoteen 1959. Ekstrapoloinnilla tarkoitetaan kuvaajan jatkamista siten, että pystytään lukemaan arvoja, joihin alkuperäinen kuvaaja ei yllä. Ekstrapoloinnissa on omat haasteensa, sillä kuvaajaa on jatkettava niin, että alkuperäisen kuvaajan suunta ja kaarevuus säilyy. Lopulliset funktiosovitteet kertoimille muodostetaan käyttäen koko käyrää vuodesta 1940 vuoteen 2012.

Tutkimusta tehtäessä naiset ja miehet jaetaan usein omiksi ryhmikseen, niin fyysisten ominaisuuksien kuin perinnöllistenkin tekijöiden perusteella. Tässäkin tutkielmassa jaottelu on hyvin perusteltua, sillä kokemuksepäisesti sukupuolen on havaittu vaikuttavan merkittävästi henkilön elinajan pituuteen. Henkivakuutuksessa miehille ja naisille käytetäänkin yleensä eri kuolevuusjakaumia (Pesonen, Soininen & Tuominen, 1999). Suomessa kuluttajille myönnettävissä vakuutuksissa ei kuitenkaan enää saa käyttää sukupuolta vakuutusmaksujen tai vakuutuksesta saatavien etuuksien laskennassa. Sukupuoleen perustuvia erojakaan maksuissa tai etuuksissa ei saa tehdä. Ennen Euroopan unionin tuomioistuimen antamaa tasa-arvodirektiiviä Suomessa käytettiin sukupuolta yhtenä maksutekijänä, sillä sukupuoli vaikuttaa esimerkiksi henkilövakuutuksilla vakuutettavien riskien todennäköisyyteen. Lakimuutos koskee kuitenkin vain kuluttajille tarjottavia vakuutuksia, joten sukupuoleen perustuvat erot ovat edelleen sallittuja muissa vakuutuksissa, jos sukupuolen käyttäminen näissä vakuutuksissa on vaikuttava tekijä tilasto- tai vakuutusmatemaattisten tietojen perusteella tehtävässä riskinarvioinnissa. Työnantajan henkilöstölleen ottamat ryhmävakuutukset kuuluvat tällaisiin vakuutuksiin, joten tutkielmassa on perusteltua käsitellä vain naisia omana populaationaan. (Finanssivalvonta, 2012.)

Kuviosta 1 nähdään, kuinka vastasyntyneiden elinajanodote on muuttunut Suomessa 41 vuodessa aikavälillä 1971–2012. Pystyakselin elinajanodote esitetään vuosissa ja vaaka-akselilla annetaan vastasyntyneiden syntymävuodet.



Kuvio 1. Vastasyntyneiden elinajanodote sukupuolittain vuosina 1971–2012. (Tilastokeskus / Kuolleet, 2014)

Kuviosta 1 näkyy ero naisten ja miesten elinajanodotteiden välillä. Suomalaisten elinaika on selkeästi pidentynyt viime vuosikymmenten aikana ja ajan myötä erot ovat kaventuneet miesten ja naisten välillä. Suomessa miesten ja naisten välinen kuolleisuusero on kuitenkin suurempi kuin useimmissa muissa läntisen Euroopan maissa. Tärkeimpänä syynä elinajan pidentymiseen voidaan pitää sepelvaltimo- ja muiden verenkiertoelinsairauksien aiheuttaman kuolleisuuden pienenemistä. Lisäksi kasvaimien sekä monien muiden tautien aiheuttama kuolleisuus on varsinkin keski-ikäisillä pienentynyt selkeästi. Toisaalta taas runsaaseen alkoholinkäyttöön liittyvät kuolemat ovat lisääntyneet huomattavasti ja työikäisillä miehillä osuus onkin yli nelinkertainen naisiin verrattuna. Sukupuolten välisistä kuolleisuuseroista työikäisillä lähes kolmannes selittyy kuitenkin tapaturmaisilla ja väkivaltaisilla kuolemilla. (Martelin, Koskinen & Valkonen, 2005.) Kuviosta nähdään myös, että nousunopeus on sekä naisten että miesten kohdalla melko tasainen.

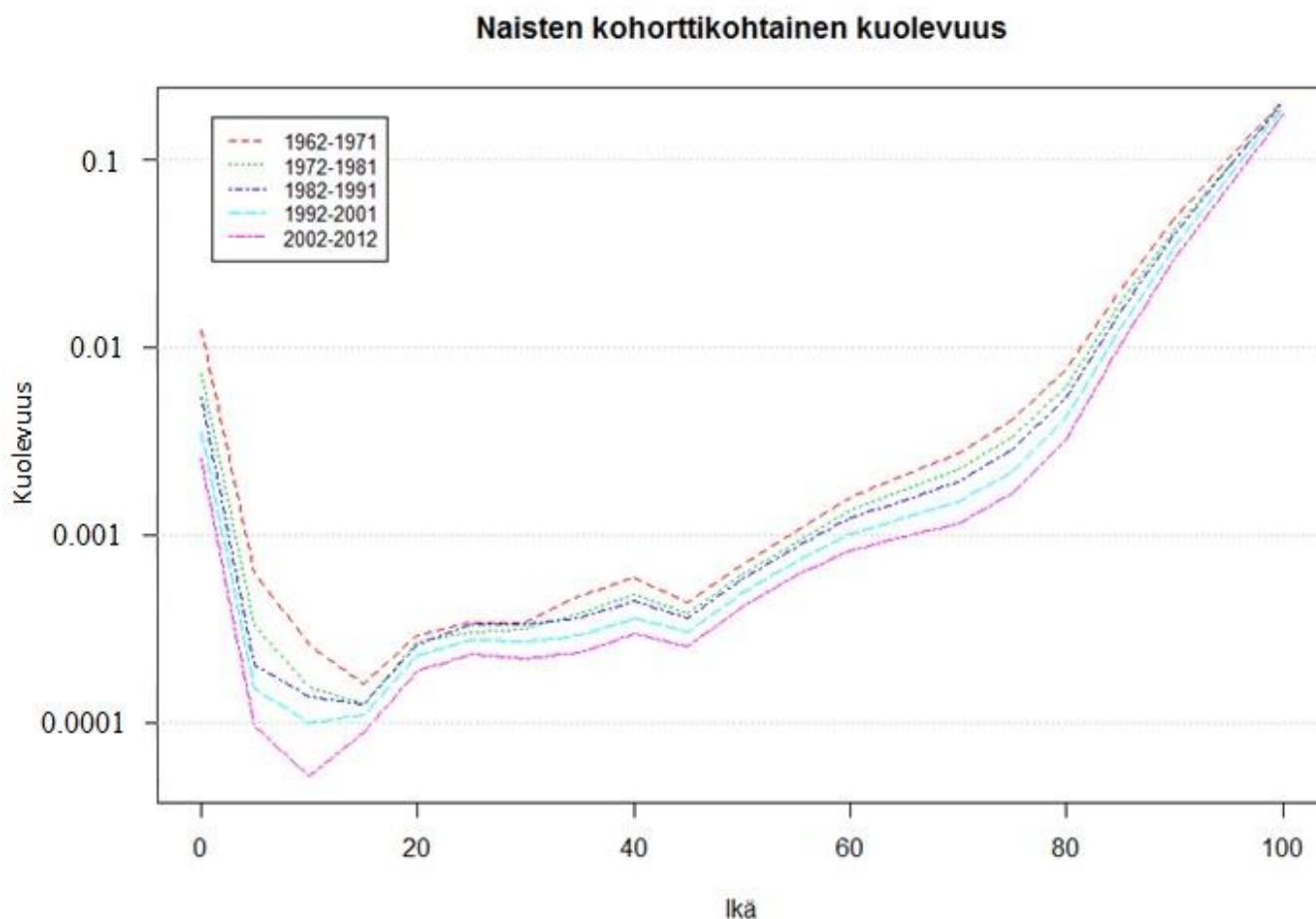
Käytännössä henkilön odotettuun elinaikaan vaikuttavat luonnollisesti myös muut tekijät kuin sukupuoli ja syntymävuosi. Tällaisia tekijöitä ovat muun muassa henkilön ikä, terveydentila, elintavat ja harrastukset. Näistä osaa on kuitenkin vaikea ottaa huomioon vakuutusta myönnettäessä ja hinnoiteltaessa. Terveydentilalla voi usein olla merkittävä vaikutus, mutta myös esimerkiksi elintapoja, kuten tupakointia voidaan käyttää maksuperusteena. Elintapojen huomioon ottaminen vakuutuksen hinnoittelussa voi kuitenkin olla hankalaa muun muassa luokittelun vaikeuden sekä elintapamuutosten takia. (Pesonen, Soininen & Tuominen, 1999.)

Voidaan miettiä, pitäisikö viimeisimmillä vuosikymmenillä olla suurempi painoarvo ennusteita tehtäessä, sillä kuten kuviosta 1 nähdään, naisten keskimääräisen elinajan pituus on kasvanut kolmena viime vuosikymmenenä hieman maltillisemmin kuin miesten. Vaikka miesten ja naisten keskimääräiset jäljellä olevat elinajanodotteet lähenevät toisiaan, ei voida kuitenkaan suoraan päätellä, että erot tasoittuisivat täysin ajan myötä, vaan naisten elinajanodote tulee todennäköisesti aina olemaan hieman korkeampi kuin miesten. (Mäkinen, 2004.)

2.1.1 Populaation henkivakuutuskuolevuus

Kuviossa 2 esitetään, kuinka kohorttikohtaiset kuolevuusfunktiot ovat muuttuneet vuosina 1962–2012. Kuolevuusfunktiot on laskettu kohorttiaineistosta naisten ryhmäeläkkeen osalta. Kuten kuviosta 2 nähdään, kuolevuus on laskenut pääsääntöisesti kaikissa ikäluokissa, mikä tukee kuviossa 1 annettuja tu-

loksia elinajanodotteen kasvusta. Kuviosta 2 voidaan myös havaita, kuinka kohorttikohtaiset kuolevuudet noudattavat tietynlaista trendiä vuosikymmenestä riippumatta.



Kuvio 2. Naisten havaittu kohorttikohtainen kuolevuus vuosina 1962–2012.

2.1.2 Lukumäärä- ja summakuolevuus

Aiemmin toteutetuissa henkivakuutuksen kuolevuustutkimuksissa on käytetty summakuolevuutta. Sen etuina on pidetty muun muassa sen kykyä ottaa huomioon henkilön varallisuus, jolla on selkeä riippuvuus henkilön sosiaaliseen asemaan. (Sirén, 2012.)

Henkivakuutusten riskiperusteanalyyseissä on volyymisuurena käytetty aiemmin niin kutsuttua riskisummaa, joka on määritelty kuoleman varalta olevan summan ja rahaston erotukseksi (SHY, 1982). Yhtiöiden käytänteissä on kuitenkin eroja. Koska vastaavasta kuolinmenosta ei ole vähennetty vapau-

tuvaan rahastoa, ovat eräät yhtiöt määritelleet riskisummaksi kuoleman varalta olevan riskisumman. Eläkevakuutuksen volyyymisuurena on käytetty rahastoa. (Sirén, 2012.)

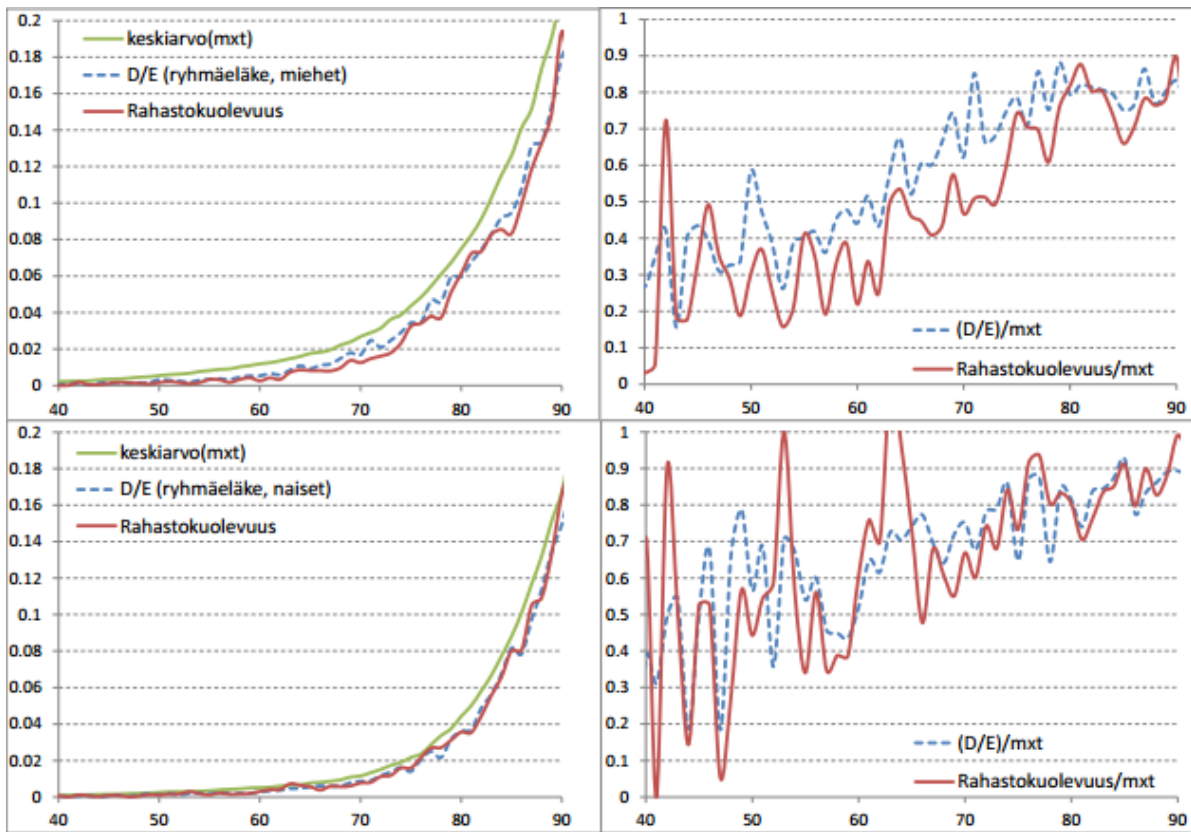
Lukumääräkuolevuus m_x^L on kuolleiden lukumäärän (D_x) suhde kuolinriskille altistumaan (E_x) ("D/E"-luku) tarkastelujaksolla t

$$m_x^L = \frac{\sum_t D_{x,t}}{\sum_t E_{x,t}}.$$

Summakuolevuus m_x^S taas määritellään seuraavasti:

$$m_x^S = \frac{\sum_t S_{x,t}^K}{\sum_t (S_{x,t} - \frac{1}{2} S_{x,t}^K)},$$

missä $S_{x,t}^K$ tarkoittaa x -ikäisenä kuolleiden riskisummaa ja rahastoja eli volyyymisuuretta vuonna t ja $S_{x,t}$ x -ikäisten volyyymisuuretta vuoden t alussa. Negatiiviset volyyymisuureet on muutettu tarkasteluissa nolliksi ja kuviossa 3 kuolevuudet on laskettu kohorteittain. Summaus on tehty ryhmäeläkkeen osalta yli vuosien 2002–2009 satunnaisvaihtelun vähentämiseksi. Lisäksi kuviossa on esitetty lukumäärä- ja summa- eli rahastokuolevuuksien suhde väestökuolevuuteen $m_{x,t}$, missä on käytetty HMD-tietokannan väestökuolevuuksien keskiarvoa vuosilta 2002–2009. (Sirén, 2012.)



Kuvio 3. Ryhmäeläkkeiden lukumäärä- ja rahastokuolevuusvertailu miehille ja naisille. Kuviossa väestökuolevuus $m_{x,t}$ on keskiarvo vuosilta 2002–2009 (HMD). Miesten ryhmäeläkkeiden kuolevuusvertailut esitetään ylemmällä ja naisten alemmalla rivillä. (Sirén, 2012.)

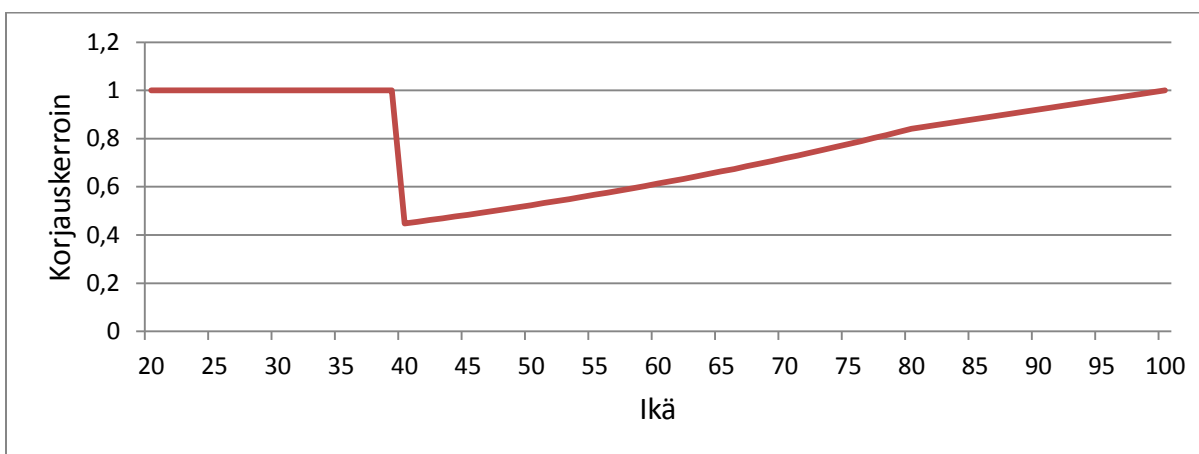
Ryhmäeläkevakuutuksessa rahasto määräytyy vakuutetun palkan mukaan, joten summakuolevuudessa ansiotason ja kuolevuuden negatiivinen korrelaatio tulee otettua huomioon. Kuviossa 3 esitetään ryhmäeläkkeen lukumäärä- sekä summakuolevuudet miehille ja naisille. Nähdään, että miesten kohdalla ero näiden kuolevuuksien välillä on selkeä. Naisilla taas kuolevuudet ovat lähempänä toisiaan. Toisin sanoen miesten kohdalla ansiotaso vaikuttaa kuolevuuteen enemmän kuin naisilla. (Sirén, 2012.) Tässä tutkielmassa korjauskertoimet määritellään lukumääräaineiston pohjalta ryhmäeläkevakuutukselle pelkästään naisten osalta eikä miesten kuolevuuksia oteta tutkielmassa huomioon.

2.1.3 Korjauserroin

Vakuutustoiminnassa otetaan huomioon valinnan vaikutus kuolevuuteen. Yleensä vapaaehtoinen eläke- tai henkivakuutus myönnetään vain, jos terveydentila on riittävän hyvä. Vakuutuksen myöntämisen jälkeen vakuutetun kuolevuus onkin tästä syystä väestökuolevuutta pienempi eli toisin sanoen vakuu-

tuskantaan valikoituu keskimääräistä terveempiä ihmisiä, joiden voi siis myös olettaa elävän pidempään. (Nyrhinen, 2012.) Henkivakuutuksen kannattavuuden kannalta on hyvin olennaista, minkälaisia riskejä yhtiö ottaa vastuulleen. Riskit voidaan esimerkiksi hinnoitella eri henkilöille eri tavoin, mutta yhtiö voi vaikuttaa valintaan myös tuote-ehdoilla sekä tuotevalikoimalla ja valitsemalla vakuutettavat henkilöt ja riskit, joille vakuutus myönnetään. Toisaalta taas henkivakuutusaineisto ei ole riittävän kattava pitkän ajan ennusteen tekemiseen. Näin ollen henkivakuutusaineiston sekä Suomen väestöaineiston avulla on muodostettu korjauskerroin, jonka avulla Suomen väestön kuolevuusennusteita voidaan korjata henkivakuutettujen kuolevuutta paremmin vastaavaksi. (Sirén, 2012.)

Ryhmäeläkkeen kohdalla havaittu kuolevuus suhteessa väestökuolevuuteen on lähes sama vanhemmissa ikäluokissa (Sirén, 2012). Kuolevuusmallia sovitettaessa ikä on rajoitettu välille 40–80, jonka jälkeen korjauskertoimet on interpoloitu välille 81–100 olettaen, että 100-vuotiaasta lähtien kerroin on jälleen yksi. Kuviossa 4 esitetään korjauskertoimet naisten tapauksessa ikävuosille 20–100. Nähdään, että korjauskertoimet nousevat lähes lineaarisesti 40 ikävuodesta aina 80 ikävuoteen saakka. Tämän jälkeen kertoimet kasvavat lineaarisesti ikävuoteen 100 asti. Ryhmäeläkevakuutuksessa alle 40-vuotiaiden henkivakuutuskuolleisuuden ja Suomen väestön kuolevuusennusteiden katsotaan vastaavan toisiaan ja näin ollen loikkaus tapahtuu vasta 40 ikävuoden kohdalla.



Kuvio 4. Korjauskertoimet ryhmäeläkevakuutukselle naisten osalta ikävuosille 20–100.

2.1.4 Ryhmäeläkevakuutus

Työnantaja voi järjestää työntekijöilleen lisäeläketurvaa ottamalla heille joko yksilöllisen eläkevakuutuksen tai ryhmähenkivakuutuksen. Eläketurvan täydentäminen lisäeläkkeen avulla on tehokas tapa parantaa henkilöstön pitkäaikaista sitoutumista (Mandatum Life, 2012). Jos lisäeläkevakuutus halutaan kohdentaa tiettyyn henkilöön, tulee tällöin kyseeseen yksilöllinen eläkevakuutus. Ryhmäeläkevakuutus taas on kollektiivinen ja sen edellytyksiin kuuluu, että sen piiriin tulevat henkilöt valitaan esimerkiksi ammattiaseman tai työtehtävän perusteella. (Eläketurvakeskus, 2014.) Aiemmin tätä etua tarjottiin lähinnä johdolle, mutta se on työeläkelain muutosten myötä laajentunut koskemaan myös muita yrityksen työntekijöitä (Mandatum Life, 2012).

Työnantajien järjestämä ryhmäeläketurva on vapaamuotoista kollektiivista lisäeläkevakuuttamista eikä vakuutusten ottaminen ole pakollista. Vakuutusten sisältö on myös jokseenkin vapaasti muotoiltavissa, mutta tavallisesti vapaamuotoinen ryhmäeläkevakuutus koskee vain vanhuuseläkettä. Siihen voidaan kuitenkin liittää myös perhe-eläke- ja työkyvyttömyysturva. (Eläketurvakeskus, 2014.)

Yleensä vapaamuotoisiin ryhmäeläkevakuutuksiin liittyy oikeus karttuneeseen lisäeläkkeeseen työsuhteen päättyessä eli niin kutsuttu vapaakirjaoikeus, joka voi esimerkiksi olla sidottu työsuhteen keston. Vapaakirjaoikeus voi myös olla osittainen, esimerkiksi 50 prosenttia karttuneesta lisäeläkkeestä. Aina oikeutta karttuneeseen lisäeläkkeeseen työsuhteen päättyessä ei kuitenkaan välttämättä ole, jolloin työntekijä menettää lisäeläke-etunsa irtisanomistilanteessa tai työpaikkaa vaihtaessaan. (Eläketurvakeskus, 2014.)

Vapaaehtoinen ryhmäeläkevakuutus voidaan järjestää henkivakuutusyhtiön, eläkesäätiön tai eläkekas-san kautta, mutta se voi perustua myös yrityksen omaan eläkesääntöön. Tällöin sitä ei kuitenkaan pidetä vakuutuksena vaan kirjanpidollisena varauksena, jossa työnantaja sitoutuu maksamaan eläkettä määritellylle henkilöryhmälle. Suomessa Finanssivalvonta valvoo vakuutuslalla toimivia eläkeyhtiöitä, säätiöitä sekä kassoja. (Eläketurvakeskus, 2014.)

Ryhmäeläkevakuutuksen tulee koskea kollektiivista, vähintään kahden henkilön suuruista henkilöryhmää. Vakuutettuna oleva ryhmä voidaan määritellä usealla eri tavalla, kuten esimerkiksi työntekijän asemaan, toimialaan tai -paikkaan, syntymäaikaan, työsuhteen alkamiseen, työnantajan järjestämään muuhun eläketurvaan tai toiminnan siirtymiseen perustuvalla jaolla. Vakuutuksesta maksettava lisäelä-

ke on joko täydentävää euromääräistä lisäeläkettä tai sen tarkoituksena on alentaa eläkeikää, mutta se voi olla myös näiden kahden etuuden yhdistelmä. (Eläketurvakeskus, 2014.)

Ryhmäeläkevakuutus voi olla etuus- tai maksuperusteinen. Etuusperusteisessa järjestelmässä työntekijälle on valmiiksi määritetty tietty etuuden taso. Maksuperusteisessa järjestelmässä taas asetetaan ainoastaan maksuosuus, ja aikanaan maksettava lisäeläke määräytyy vuosien mittaan kertyneen vakuutus- säästön mukaan. Tänä päivänä solmitaan lähes ainoastaan maksuperusteisia ryhmäeläkevakuutuksia. (Eläketurvakeskus, 2014.)

Tällä hetkellä lähes joka neljäs yksityisen sektorin työeläkkeensaaja saa myös lisäeläkettä, mutta työnantajan lisäeläketurvan merkitys on euromääräisesti kuitenkin vähäinen. Ryhmäeläkkeet muodostavat maksussa olevista eläkkeistä vain alle neljä prosenttia. (Eläketurvakeskus, 2014.)

2.2 Korkoaineisto

Toinen tutkielmassa käytettävä aineisto sisältää The Federal Reserven 31 vuoden ajalta mitatun korkoaineiston, tammikuusta 1982 joulukuuhun 2012. Korot ovat markkinatuottoja Yhdysvaltojen liikkeelle laskemista joukkovelkakirjoista eri maturiteeteilla. Maturiteetit vaihtelevat kolmesta kuukaudesta kymmeneen vuoteen, joten aineiston yhtenäistämiseksi maturiteetit annetaan kuukausina, jolloin kaikki kahdeksan maturiteettia ovat 3, 6, 12, 24, 36, 60, 84 ja 120. Arvopaperit on noteerattu investointien perusteella ja luvut ovat kuukausittaisia keskiarvoja.

Markkinakorkojen muutokset näkyvät koko sijoitussalkussa. Korkosijoituksissa arvopaperin käypä arvo määräytyy korkopaperin kupongista ja sen koron myötä, jolla arvopaperista syntyvät tulevat kasviritrat diskontataan. (Saraste, 2009.) Diskonttauksessa käytettävän diskonttokoron avulla eri ajankoh- tien rahamäärät muunnetaan nykyarvoksi PV (Present value), mikä mahdollistaa eri ajankohtina toteu- tujen rahavirtojen vertailun. Diskonttaustekijä $v_t(\tau)$ määritellään

$$(2.1) \quad v_t(\tau) = \frac{1}{(1 + y_t(\tau))^\tau},$$

missä t on tarkasteluajankohta ja $y_t(\tau)$ on korkokanta maturiteetilla τ (Vedenjuoksu, 2011). Tutkiel- massa diskonttauksessa käytetään simuloitua korkokäyrää, ja koska korkokäyrä ei ole tasainen on jo- kainen tuleva rahamäärä diskontattava käyttämällä senhetkistä korkokantaa.

Tutkielmassa muodostettavan dynaamisen Nelson-Siegel-korkomallin perusteella tehdään ennuste vuoden päähän simuloimalla. Ennusteen avulla lasketaan elinikäisen eläkkeen pääoma-arvoja valitulle annuiteettikohortille. SCR-pääomavaatimuksessa vakuutusyhtiön oman varallisuuden OF muutokselle vuoden aikana lasketaan 0,5 % kvantiilin vastaluku eli niin kutsuttu VaR-luku. Omalla varallisuudella $OF(t)$ tarkoitetaan vakuutusyhtiön markkina-arvoisen taseen omaisuuden ja velkojen erotusta hetkellä t , jota tutkielmassa kuvataan obligaatiosijoitusten arvon ja eläkevakuutusten pääoma-arvon erotuksella.

3 TUTKIMUKSESSA KÄYTETYT MENETELMÄT

Tämä luku sisältää tutkimuksen aineistojen analysoinnissa käytetyt menetelmät. Kuolevuusaineiston funktiosovitteen tekemiseen käytettyä regressioanalyysia käsitellään ensin, jonka jälkeen siirrytään korkomallien muodostamisessa käytettyyn aikasarja-analyysiin.

3.1 Regressioanalyysi

Luvussa esitellään pintapuolisesti lineaarista regressiota keskittyen pääasiassa epälineaarisen regressioanalyysin teoriaan. Molempia luvussa käsiteltyjä malleja käytetään myöhemmin muodostettaessa funktiosovite kuolevuuksille sekä korkomalleja tarkasteltaessa.

3.1.1 Lineaarinen regressio

Regressiofunktio $\mu_y(x_1, \dots, x_k)$ kuvaa jonkin riippuvan eli selitettävän muuttujan y vaihtelua muiden riippumattomien eli selittävien muuttujien x_1, \dots, x_k avulla (Graybill & Iyer, 1994). Lineaarinen regressiomalli on muotoa

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + \varepsilon,$$

missä y on satunnaismuuttuja, muuttujat x_1, \dots, x_k ei-satunnaisia muuttujia, ε virhetermi oletuksella $E(\varepsilon) = 0$ sekä parametrit β_0 ja β_1, \dots, β_k aineiston perusteella estimoitavat mallin tuntemattomat parametrit (Seber & Wild, 1989). Lineaarisen regressioanalyysin oletus on, että kahden muuttujan välinen riippuvuus on lineaarista. Jos ehto lineaarisuudesta ei kuitenkaan täysin täyty, saattaa lineaarinen malli joskus olla sopiva havainnollistamaan muuttujien välistä riippuvuutta (Graybill & Iyer, 1994). Seuraavassa esimerkkejä lineaarisista regressiofunktioista:

$$(3.1) \quad \mu_y(x) = \beta_0,$$

$$(3.2) \quad \mu_y(x) = \beta_0 + \beta_1 x,$$

$$(3.3) \quad \mu_y(x) = \beta_0 + \beta_1 x + \beta_2 x^2,$$

missä x on ennustemuuttuja sekä β_0, β_1 ja β_2 ovat tuntemattomia parametreja. Mallit (3.1) – (3.3) ovat lineaarisia regressiomalleja, sillä kussakin tapauksessa regressiofunktio on parametrien suhteen lineaar-

rinen. Mallissa (3.3) nähdään toisen asteen termi mutta se on silti lineaarinen. Tällaista lineaarisen mallin laajentamista sovittamaan epälineaarisia suhteita kutsutaan polynomiregressioksi (Hastie, James, Tibshirani & Witten, 2013).

3.1.2 Epälineaarinen regressio

Kaikkia Makeham-funktiosovitteen kertoimia ei pystytä estimoimaan lineaarisen regressiomallin avulla, vaan osa estimoidaan käyttäen hyödyksi epälineaarista regressiota. Epälineaarisessa regressiomallissa regressiofunktio on epälineaarinen ainakin yhden parametrin suhteen. Mallia kuvataan yhtälöllä

$$(3.4) \quad y_i = f(\theta; x_i) + \varepsilon_i, \quad i = 1, \dots, n,$$

missä funktio y_i on tuntematonta parametrivektoria θ vaille tunnettu ja stokastisen virhetermin ε_i oletetaan yleisesti noudattavan normaalijakaumaa odotusarvolla 0 ja varianssilla σ^2 . Parametrien estimointiin voidaan käyttää useita eri menetelmiä, kuten pienimmän neliösumman tai suurimman uskottavuuden menetelmää tai robusteja tai bayesiläisiä estimointimenetelmiä. (Seber & Wild, 1989). Tässä tutkielmassa keskitytään kuitenkin vain pienimmän neliösumman menetelmään, sillä kun funktiosovitetta tehdään ja malleja tarkastellaan, sitä käytetään parametrien estimointiin. Esimerkkejä epälineaarisista malleista ovat

$$(3.5) \quad \mu_y(x_1) = \beta_1 e^{\beta_2 x_1},$$

$$(3.6) \quad \mu_y(x_1) = \beta_0 + \beta_1 e^{\beta_2 x_1},$$

$$(3.7) \quad \mu_y(x_1, x_2) = \beta_0 + \beta_1 e^{\beta_2 x_1} + \beta_3 e^{\beta_4 x_2},$$

missä x_1 ja x_2 ovat ennustemuuttujia ja β_0, \dots, β_4 tuntemattomia parametreja. Mallit (3.5) – (3.7) ovat ainakin yhden parametrin suhteen epälineaarisia, mikä erottaa ne lineaarisista regressiomalleista.

Jotta pystytään ymmärtämään parametrien estimointia pienimmän neliösumman menetelmällä, käytetään apuna regressiomallia (3.4). Kun virhetermi oletetaan normaalijakautuneeksi, epälineaarisen regressiomallin uskottavuusfunktio voidaan antaa muodossa

$$L(\theta; \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp \left\{ -\frac{\sum_{i=1}^n [y_i - f(\theta, x_i)]^2}{2\sigma^2} \right\}.$$

Suurimman uskottavuuden estimaattori $\hat{\theta}$ on θ :n arvo, joka minimoi neliöjäännösten summan

$$S(\theta) = \sum_{i=1}^n [y_i - f(\theta; x_i)]^2.$$

Muodostetaan seuraavaksi yhtälöstä osittaisderivaatta

$$\frac{\partial S(\theta)}{\partial \theta} = -2 \sum_{i=1}^n [y_i - f(\theta; x_i)] \frac{\partial f(\theta; x_i)}{\partial \theta}$$

ja asettamalla osittaisderivaatta nolaksi voidaan ratkaista yhtälön regressiokertoimet. (Fox, 2002.)

3.2 Aikasarja-analyysi

Kun jonkin muuttujan x arvoja mitataan peräkkäisinä ajankohtina t , syntyy lukujono x_t , jota kutsutaan aikasarjaksi. Matemaattisena mallina toimii stokastinen prosessi $\{X_t, t \in T\}$, missä indeksijoukko T voi olla joko diskreetti tai jatkuva. Esimerkkejä aikasarjoista ovat muun muassa talouden kehittymisen tai teollisuuden prosessin seuraaminen. Usein aikasarjoista on saatavilla yksi realisaatio tietyllä aikavälillä, jota analysoimalla pyritään löytämään stokastinen prosessi, joka voisi tuottaa kyseisen sarjan. Tämä voisi auttaa ensinnäkin ymmärtämään jotain tutkittavan ilmiön luonteesta ja toiseksi ennustamaan sarjan tulevia arvoja. Menetelmää sovelletaan myös prosessien kontrolloinnissa, esimerkiksi teollisuudessa. Lisäksi yhden sarjan vaihtelua voidaan pyrkiä selittämään muiden sarjojen vaihtelulla, jolloin sivutaan samankaltaisia ongelmia kuin regressioanalyysissä. (Brockwell & Davis, 2002.)

3.2.1 Stationaarinen prosessi

Aikasarjaa, jossa ei esiinny kausivaihtelua, trendiä tai varianssin muutosta ajan suhteen ja jonka autokorrelaatorakenne pysyy samanlaisena ajan kuluessa, kutsutaan stationaariseksi sarjaksi (Cryer & Chan, 2008). Aikasarja pyritään yleensä erilaisilla muunnoksilla saattamaan sellaiseen muotoon, että sen nähdään olevan realisaatio stationaarisesta prosessista. Yleensä stationaarista prosessia voidaan tilastollisesti mallintaa niin kutsutun ARMA-prosessin avulla. Tämän jälkeen alkuperäinen aikasarja voidaan kuvata käyttäen stationaarista aikasarjaa sekä niitä muunnoksia, joiden avulla siitä saadaan alkuperäinen sarja. (Brockwell & Davis, 2002.)

Jos peräkkäisten havaintojen yhteisjakauma ei muutu, kun siirrytään ajassa eteenpäin, prosessin $\{X_t\}$ sanotaan olevan vahvasti stationaarinen. Puhuttaessa stationaarisuudesta tarkoitetaan kuitenkin yleensä niin kutsuttua heikkoa stationaarisuutta. Jos peräkkäisten havaintojen kovarianssimatriisi eikä prosessin odotusarvo muutu, kun siirrytään ajassa eteenpäin, prosessin $\{X_t\}$ sanotaan olevan heikosti stationaarinen. Kovarianssifunktio voidaan ilmaista $\gamma_X(t+h, t) = \gamma_X(h, 0) = \gamma_X(h)$, jos prosessi on stationaarinen. Määritellään odotusarvofunktio $\mu_X(t) = E(X_t)$ ja autokovarianssifunktio $\gamma_X(r, s) = \text{Cov}(X_r, X_s)$ kaikille kokonaisluvulle r ja s . Määriteltyjen funktioiden avulla prosessi on heikosti stationaarinen, jos odotusarvofunktio on ajasta t riippumaton vakio eikä kovarianssifunktio riipu ajasta t millään kokonaisluvulla h . Stationaariselle prosessille voidaan lisäksi määritellä autokorrelaatiofunktio $\rho_X(h) = \gamma_X(h)/\gamma_X(0) = \text{Cor}(X_{t+h}, X_t)$. (Brockwell & Davis, 2002.) Aikasarjan tulee olla vähintään heikosti stationaarinen, jotta sitä voidaan mallintaa liukuvan keskiarvon tai autoregressiivisen prosessin avulla (Cryer & Chan, 2008). Seuraavaksi määritellään erilaisia aikasarja-analyysin prosesseja.

Jos $\{X_t\}$ on jono riippumattomia ja samoin jakautuneita satunnaismuuttujia, joiden odotusarvo on 0, kutsutaan jonoa *IID-kohinaksi*. Tällöin autokovarianssifunktio on

$$\gamma_X(t+h, t) = \begin{cases} \sigma^2, & \text{kun } h = 0, \\ 0, & \text{kun } h \neq 0, \end{cases}$$

mikä ei riipu ajasta t . Jonoa merkitään $\{X_t\} \sim \text{IID}(0, \sigma^2)$ ja sitä voidaan pitää ehkä yksinkertaisimpana esimerkkinä vahvasta stationaarisuudesta. (Brockwell & Davis, 2002.)

Kuten edellä, jonoa sanotaan *valkoiseksi kohinaksi*, jos $\{X_t\}$ on jono korreloimattomia satunnaismuuttujia odotusarvolla 0 ja varianssilla σ^2 . Jonoa merkitään $\{X_t\} \sim \text{WN}(0, \sigma^2)$ ja se on esimerkki heikosti stationaarisesta prosessista, sillä määritelmän mukaan $\mu_X(t) = EX_t = 0$ ja

$$\gamma_X(t+h, t) = \begin{cases} \sigma^2, & \text{kun } h = 0, \\ 0, & \text{kun } h \neq 0. \end{cases}$$

Nähdään, etteivät odotusarvo- ja autokovarianssifunktiot riipu ajasta t . (Brockwell & Davis, 2002.) Jotta korrelaation mallintaminen olisi järkevää aloittaa, aikasarja ei saa olla valkoista kohinaa. Toisaalta kuitenkin, jotta sarjan mallia voidaan pitää riittävänä, tulisi mallin jäännössarjan muistuttaa valkoista kohinaa.

Jos jono $\{Z_t\} \sim \text{WN}(0, \sigma^2)$ ja määritellään $X_t = Z_t + \theta Z_{t-1}$, missä $t = 0, \pm 1, \dots$ ja θ on reaaliluku, sanotaan prosessin $\{X_t\}$ noudattavan ensimmäisen asteen liukuvan keskiarvon prosessia eli *MA(1)-prosessia*. Tällöin $\mu_X(t) = 0$ ja $\text{E}X_t^2 = \sigma^2(1 + \theta^2) < \infty$. Lisäksi

$$\gamma_X(t+h, t) = \begin{cases} \sigma^2(1 + \theta^2), & \text{kun } h = 0, \\ \sigma^2\theta, & \text{kun } h = \pm 1, \\ 0, & \text{kun } |h| > 1. \end{cases}$$

Koska $\mu_X(t)$ eikä $\gamma_X(t+h, t)$ riipu ajasta t , on kyseessä stationaarinen prosessi. Autokorrelaatiofunktio prosessille $\{X_t\}$ on

$$\rho_X(t+h, t) = \begin{cases} 1, & \text{kun } h = 0, \\ \theta/(1 + \theta^2), & \text{kun } h = \pm 1, \\ 0, & \text{kun } |h| > 1. \end{cases}$$

(Brockwell & Davis, 2002.)

Prosessin $\{X_t\}$ sanotaan noudattavan ensimmäisen asteen autoregressiivistä prosessia, eli *AR(1)-prosessia*, jos $\{X_t\}$ on stationaarinen aikasarja, joka toteuttaa yhtälöt

$$(3.8) \quad X_t = \phi X_{t-1} + Z_t, \quad t = 0, \pm 1, \dots,$$

missä $\{Z_t\} \sim \text{WN}(0, \sigma^2)$, $|\phi| < 1$ ja Z_t on korreloimaton satunnaismuuttujien X_s kanssa, kun $s < t$. Kun lasketaan odotusarvo yhtälön (3.8) eri puolista, saadaan odotusarvoksi $\text{E}X_t = 0$. Jotta voidaan määrittää autokovarianssifunktio, kerrotaan yhtälö puolittain havainnolla X_{t-h} , missä $h > 0$, ja lasketaan odotusarvo puolittain

$$\begin{aligned} \text{E}(X_t X_{t-h}) &= \phi \text{E}(X_{t-1} X_{t-h}) + \text{E}(Z_t X_{t-h}) \\ \Leftrightarrow \text{Cov}(X_t, X_{t-h}) &= \phi \text{Cov}(X_{t-1}, X_{t-h}) + 0 \\ \Leftrightarrow \gamma_X(h) &= \phi \gamma_X(h-1). \end{aligned}$$

Positiivisille ja negatiivisille viiveille h soveltuva autokovarianssin kaava on $\gamma_X(h) = \phi^{|h|} \gamma_X(0)$, sillä

$$\gamma_X(h) = \text{Cov}(X_{t+h}, X_t) = \text{Cov}(X_t, X_{t+h}) = \gamma_X(-h).$$

Autokorrelaatio taas on

$$\rho_X(h) = \gamma_X(h)/\gamma_X(0) = \phi^{|h|}, \quad h = 0, \pm 1, \dots$$

Jotta voidaan määrittää, mikä $\gamma_X(0)$ on, todetaan aluksi, että

$$\text{Cov}(X_t, Z_t) = \text{Cov}(\phi X_{t-1} + Z_t, Z_t) = \phi \text{Cov}(X_{t-1}, Z_t) + \text{Cov}(Z_t, Z_t) = \text{Cov}(Z_t, Z_t) = \sigma^2,$$

koska oletuksen mukaan $\text{Cov}(X_{t-1}, Z_t) = 0$. Näin ollen

$$\gamma_X(0) = \text{Cov}(X_t, X_t) = \text{Cov}(X_t, \phi X_{t-1} + Z_t) = \phi \gamma_X(1) + \sigma^2 = \phi^2 \gamma_X(0) + \sigma^2,$$

josta pystytään ratkaisemaan, että $\gamma_X(0) = \sigma^2/(1 - \phi^2)$. (Brockwell & Davis, 2002.)

Prosessia kutsutaan *satunnaiskävelyksi*, jos $\{X_t, t = 1, 2, \dots\}$ on jono riippumattomia ja samoin jakautuneita satunnaismuuttujia, joille $E(X_t) = \mu$ ja $\text{Var}(X_t) = \sigma^2$. Lisäksi kun $t \geq 1$ tarkastellaan satunnaisprosessia $\{Y_t, t = 1, 2, \dots\}$, missä $Y_0 = 0$ ja $Y_t = X_1 + X_2 + \dots + X_t$. Jos $\mu = 0$, on kyseessä symmetrinen satunnaiskävely ja jos X_t voi saada arvokseen joko -1 tai 1 , on kyseessä yksinkertainen satunnaiskävely. Prosessin odotusarvofunktio on $\mu_Y(t) = E(X_1 + X_2 + \dots + X_t) = t\mu$ ja $E(Y_t^2) = t\sigma^2 < \infty \forall t$ ja kun $h \geq 0$

$$\begin{aligned} \gamma_Y(t+h, t) &= \text{Cov}(Y_{t+h}, Y_t) \\ &= \text{Cov}(Y_t + X_{t+1} + \dots + X_{t+h}, Y_t) \\ &= \text{Cov}(Y_t, Y_t) \\ &= t\sigma^2. \end{aligned}$$

Vaikka odotusarvo olisi nolla, ei prosessi ole stationaarinen, sillä $\gamma_Y(t+h, t)$ riippuu ajasta t . Vaikkei symmetrinen satunnaiskävely ole stationaarinen, se muistuttaa AR(1)-prosessia, kun ϕ on lähellä ykköstä, jolloin $\hat{\rho}(h)$ vähenee hitaasti ykkösestä, kun $|h|$ kasvaa. (Brockwell & Davis, 2002.)

3.2.2 Otosauto- ja otos-osittaisautokorrelaatiofunktio

Yleensä aikasarjan autokovarianssi- ja autokorrelaatiofunktioita ei käytännössä tunneta, vaan ne estimoidaan otoksen perusteella. Jotta voidaan määrittää riippuvuuden aste aineistossa ja valita sitä kuvaava malli, käytetään otosautokorrelaatiofunktioita. Jos $\{x_1, x_2, \dots, x_n\}$ on otos jostain stationaarisesta aikasarjasta, määritellään tällöin otoskeskiarvo

$$\bar{x} = \frac{1}{n} \sum_{t=1}^n x_t,$$

otosautokovarianssifunktio

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} (x_{t+|h|} - \bar{x})(x_t - \bar{x}), \quad -n < h < n$$

ja otosautokorrelaatiofunktio

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}, \quad -n < h < n,$$

missä $\hat{\gamma}(h) = \text{Cov}(x_t, x_{t-h})$ on siis prosessin autokovarianssi viiveellä h ja $\hat{\gamma}(0) = \text{Var}(x_t) = \sigma^2$ prosessin varianssi. Otosautokorrelaatiofunktioista voidaan tehdä päätelmiä aikasarjan generoineen satunnaisprosessin autokorrelaatiofunktion rakenteesta, mikä saattaa auttaa sopivan mallin löytymisessä. (Brockwell & Davis, 2002.) Otos-autokorrelaatiokerroin mittaa h aikavälin etäisyydellä olevien havaintojen välistä lineaarista riippuvuutta.

Otosautokorrelaatio voidaan laskea vaikka sarja ei olisikaan stationaarinen. Tällöin siitä voidaan mahdollisesti havaita ei-stationaarisuus. Jos esimerkiksi sarjassa esiintyy lineaarinen trendi, otosautokorrelaatiofunktio pysyy pitkään positiivisena, kun h kasvaa, mikä johtuu siitä, että sarjan alkupää on keskiarvon toisella ja loppupää toisella puolella. Tällöin otosautokovarianssin lausekkeessa tekijät $x_{t+|h|} - \bar{x}$ ja $x_t - \bar{x}$ ovat samanmerkkisiä eivätkä niiden tulot siis kumoa toisiaan. Lisäksi, jos sarjassa on havaittavissa säännöllistä jaksollista vaihtelua, on otosautokorrelaatiofunktiossa myös havaittavissa samankaltaista vaihtelua samalla jaksolla. (Brockwell & Davis, 2002.)

ARMA-prosessin $\{X_t\}$ osittaisautokorrelaatiofunktio $\alpha(h)$ määritellään yhtälöillä

$$\begin{aligned} \alpha(0) &= 1, \\ \alpha(h) &= \phi_{hh}, \quad h \geq 1, \end{aligned}$$

missä ϕ_{hh} on vektorin $\phi_h = \Gamma_h^{-1} \gamma_h$ viimeinen komponentti, missä $\Gamma_h = [\gamma(i-j)]_{i,j=1}^h$. Lisäksi asetetaan $\gamma_h = [\gamma(1), \gamma(2), \dots, \gamma(h)]'$. Otos-osittaisautokorrelaatiofunktio voidaan vastaavasti määritellä havainnoille x_1, x_2, \dots, x_n yhtälöillä

$$\begin{aligned}\hat{\alpha}(0) &= 1, \\ \hat{\alpha}(h) &= \hat{\phi}_{hh}, \quad h \geq 1,\end{aligned}$$

missä $\hat{\phi}_{hh}$ on vektorin $\hat{\phi}_h = \hat{\Gamma}_h^{-1}\hat{\gamma}_h$ viimeinen komponentti. (Brockwell & Davis, 2002.) Osittaisautokorrelaatiokerroin voidaan määritellä myös korrelaationa. Prosessin h osittaisautokorrelaatiokerroin

$$\phi_h = \text{Cor}(x_t, x_{t-h} \mid x_{t-1}, \dots, x_{t-h+1}), \quad h = 0, 1, 2, \dots$$

on satunnaismuuttujien x_t ja x_{t-h} ehdollinen korrelaatio, kun ehtomuuttujina toimivat muuttujat $x_{t-1}, \dots, x_{t-h+1}$. Toisin sanoen osittaisautokorrelaatiokerroin mittaa h aikavälin etäisyydellä olevien havaintojen riippuvuutta, kun väliin jäävien havaintojen vaikutus on poistettu. (Mellin, 2010.)

3.2.3 Jäännössarjan valkoisuuden testaaminen

Aikasarjan mallintamisessa tavoitteena on saada stationaarinen jäännössarja. Jos jäännössarja muistuttaa valkoista kohinaa, sarjan mallintaminen voidaan lopettaa ja ennusteet voidaan laatia kausivaihtelun ja trendin perusteella. Jos taas jäännöstermeillä esiintyy korrelaatiota, voidaan korrelaatorakennetta mallintaa ja käyttää hyväksi ennustamisessa. Sen vuoksi onkin tärkeää testata, vastaako jäännössarja valkoisen kohinan prosessia. Esimerkiksi jos aikasarja noudattaa $\text{IID}(0, \sigma^2)$ -prosessia, otosautokorrelaatiokertoimet ovat riippumattomia ja likimain $N(0, 1/n)$ -jakautuneita. Tämä tarkoittaa sitä, että noin 95 % otosautokorrelaatiokertoimista on rajojen $\pm 1.96/\sqrt{n}$ sisällä. Esimerkiksi jos autokorrelaatiokertoimet on laskettu viiveeseen 60 asti ja useampi kuin kuusi on näiden rajojen ulkopuolella, voidaan 5 % riskitasolla hylätä hypoteesi valkoisen kohinan prosessista. Hypoteesi voidaan hylätä myös, jos jokin otosautokorrelaatiokertoimista on huomattavasti rajojen ulkopuolella, vaikka muuten kertoimet pysyisivätkin rajojen sisällä. (Brockwell & Davis, 2002.)

Aikasarjojen ja niiden jäännössarjojen valkoisuutta voidaan testata useilla eri menetelmillä. Näistä muutamia ovat Box-Ljung-testi, käännepistetestti sekä laajennettu Dickey-Fuller-testi. Näitä menetelmiä tullaan käyttämään myöhemmin tutkielman sarjojen testaamisessa. Jos jäännössarjan havaintojen normaalijakautuneisuutta halutaan tutkia sen valkoisuuden lisäksi, voidaan käyttää kvantiili-kvantiili-kuviota sekä Jarque-Bera-testiä, joka perustuu huipukkuuden ja vinouden laskemiseen. Määritellään seuraavaksi tutkielmassa käytettävät menetelmät, joilla voidaan testata aikasarjojen ja niiden jäännössarjojen valkoisuutta.

Otosautokorrelaation tutkimiseen voidaan käyttää Box-Ljung-testisuuretta

$$Q_{LB} = n(n+2) \sum_{j=1}^h \hat{\rho}^2(j)/(n-j),$$

missä $\hat{\rho}(j)$ on j . otosautokorrelaatio. Testisuure noudattaa likimain $\chi^2(h)$ -jakaumaa, kun aikasarjan voidaan katsoa olevan valkoista kohinaa eli kun testin nollahypoteesi pitää paikkansa. (Brockwell & Davis, 2002.)

Jos y_1, y_2, \dots, y_n on jono havaintoja, sanotaan, että jos $y_{i-1} < y_i$ ja $y_i > y_{i+1}$ tai $y_{i-1} > y_i$ ja $y_i < y_{i+1}$, on ajanhetkellä i käännepiste. Käännepistetestillä testataan, esiintyykö peräkkäisten havaintojen välillä autokorrelaatiota. IID-sarjassa käännepisteen todennäköisyys pisteessä i on $2/3$. Kun käännepisteiden määrä n -pituudessa sarjassa on T , saadaan sarjan odotusarvoksi

$$\mu_T = E(T) = 2(n-2)/3.$$

Voidaan myös osoittaa, että

$$\sigma_T^2 = \text{Var}(T) = (16n-29)/90.$$

Jos $T - \mu_T$ on merkittävästi nollaa pienempi, on se merkki positiivisesta autokorrelaatiosta. Jos taas $T - \mu_T$ on merkittävästi nollaa suurempi, voidaan päätellä, että sarja vaihtaa suuntaa nopeammin kuin IID-prosessin tapauksessa voitaisiin olettaa. Kun n on suuri ja aikasarja noudattaa IID-prosessia, niin $T \sim N(\mu_T, \sigma_T^2)$. (Brockwell & Davis, 2002.)

Epästationaarista aikasarjaa ei aina voida saattaa stationaariseksi vähentämällä siitä kausikomponentti tai deterministinen termi. Sarja on epästationaarinen, jos autoregressiivisellä polynomilla $\phi(z)$ on yksikin nollakohta yksikköympyrällä. Jos polynomilla on nolla kohdassa $z = 1$, ilmenee se otosautokorrelaatiofunktion vähenemisenä hitaasti ykkösestä. Aikasarja näyttää satunnaiskävelyä, jolloin sen varianssi kasvaa ajan myötä eikä sillä näytä olevan kiinteää keskiarvoa. Laajennetulla Dickey-Fuller-testillä voidaan testata yksikköjuuren olemassaoloa. AR(1)-prosessin tapauksessa testisuuretta käytetään hypoteesin $H_0: \phi_1 = 1$ testaamiseen hypoteesia $H_1: \phi_1 < 1$ vastaan ja jos nollahypoteesi hylätään, tarkoittaa se sitä, ettei sarjaa tarvitse differoida stationaarisuuden saavuttamiseksi. Astetta p olevan autoregressiivisen prosessin eli AR(p)-prosessin yhtälö on

$$X_t - \mu = \phi_1(X_{t-1} - \mu) + \dots + \phi_p(X_{t-p} - \mu) + Z_t, \quad \{Z_t\} \sim \text{WN}(0, \sigma^2),$$

joka voidaan myös kirjoittaa muodossa

$$\nabla X_t = \phi_0^* + \phi_1^* X_{t-1} + \phi_2^* X_{t-2} + \dots + \phi_p^* X_{t-p+1} + Z_t, \quad \{Z_t\} \sim \text{WN}(0, \sigma^2),$$

missä $\phi_0^* = \mu(1 - \phi_1 - \dots - \phi_p)$, $\phi_1^* = \sum_{i=1}^p \phi_i - 1$ ja $\phi_j^* = -\sum_{i=j}^p \phi_i$, $j = 2, \dots, p$. Differoitu sarja $\{\nabla X_t\}$ on AR(p-1)-prosessi ja $\phi_1^* = 0$, jos autoregressiivisellä polynomilla $\phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p$ on nollakohta $z = 1$. Nyt siis $H_0: \phi_1^* = 0$ on yhtäpitävää sen kanssa, että autoregressiivisellä polynomilla $\phi(z)$ on nolla kohdassa $z = 1$. (Brockwell & Davis, 2002.)

Jos X_1, X_2, \dots, X_n on satunnaisotos jakaumasta $N(0, 1)$ ja Y_1, Y_2, \dots, Y_n jakaumasta $N(\mu, \sigma^2)$ niin otosten järjestystunnuslukuja merkitään $X_{(i)}$ ja $Y_{(i)}$, $i = 1, 2, \dots, n$, jolloin $X_{(1)} < X_{(2)} < \dots < X_{(n)}$ ja vastaavasti $Y_{(1)} < Y_{(2)} < \dots < Y_{(n)}$. Tällöin

$$E(Y_{(j)}) = \mu + \sigma m_j,$$

missä $m_j = E(X_{(j)})$, $j = 1, 2, \dots, n$. Tästä johtuen piirrettäessä parien $(m_1, Y_{(1)}), \dots, (m_n, Y_{(n)})$ muodostama pisteparvi, pisteiden tulisi sijaita suurin piirtein suoralla. Jos sarjasta piirrettyssä kuviossa havainnot poikkeavat suoralta, tiedetään, etteivät havainnot noudata normaalijakaumaa. Kuviota kutsutaan kvantiili-kvantiili-kuvioksi.

4 TUTKIMUKSESSA KÄYTETYT MALLIT

Tässä luvussa esitellään erilaisia kuolevuus- ja korkomalleja. Aluksi käydään läpi erilaisia kuolevuusmalleja, minkä jälkeen siirrytään korkomalleihin.

4.1 Kuolevuusmallit

Kuolevuutta pystytään mallintamaan useilla menetelmillä, joista Suomessa käytössä ovat sekä Gompertz että Makeham. Ensiksi mainittu kuolevuusmalli kehitettiin vuonna 1825 ja on muotoa

$$\mu_{xt} = b \cdot e^{cx} = e^{\ln(b)+cx},$$

missä b ja c ovat aineistosta estimoitavat vakiot, x on ikä ja t viittaa kohorttiin. Mallissa kuolevuus kasvaa verrannollisesti kyseisen iän kuolevuuteen ja onkin Suomessa yleisesti käytössä lakisääteisessä eläkevakuutuksessa. (Nyrhinen, 2012.)

Makeham laajensi Gompertzin mallia 35 vuotta myöhemmin lisäämällä mukaan vakiotermiä kuvaamaan tapaturmaisen kuoleman mahdollisuutta. Makehamin kuolevuusmalli on muotoa

$$(4.1) \quad \mu_{xt} = a + b \cdot e^{cx} = a + e^{\ln(b)+cx},$$

missä b ja c ovat aineistosta estimoitavat vakiot ja parametri a edellä mainittu vakiotermi, joka on oletettu iästä riippumattomaksi. (Nyrhinen, 2012.) Kuten edellä, x tarkoittaa ikää ja t kohorttia. Malli on yleisesti vapaaehtoisessa henkivakuutuksessa käytettävä kuolevuusmalli. (Doray, 2008.)

Muita malleja ovat muun muassa De Moivre sekä logistinen kuolevuusmalli. Malleista ensimmäinen kehitettiin jo vuonna 1724 ja vaikka De Moivren yksiparametrinen malli ottaa huomioon kuolevuuden suhteellisen kasvun iän myötä, se ei muutoin sovi yhteen havaitun kuolevuuden kanssa. Mallissa elinajan ylärajana toimii 86 vuotta ja se on muotoa

$$\mu_x = \frac{1}{86 - x}, 0 \leq x \leq 86.$$

De Moivren malli ei siis sovi sellaisenaan yhteen nykyisten elinajanodotteiden kanssa, sillä niiden on havaittu kasvavan jatkuvasti tasaista, mutta varmaa vauhtia eikä takarajaa näin ollen voida asettaa. (Forfar, 2004.)

Logistisessa kuolevuusmallissa kuolevuus ei kasva suuremmilla elinaikamuuttujan x arvoilla yhtä voimakkaasti kuin Makehamin kehittämässä mallissa. Tämän vuoksi, kun tarkastelun kohteena ovat korkeat ikävuodet, logistisen kuolevuuden on todettu kuvaavan oikein valituilla parametreilla havaittua kuolevuutta Makehamin mallia paremmin. Beardin (1963) edeltävistä logistisista malleista kehittämä kuolevuusmalli on muotoa

$$\mu_x = \frac{b \cdot e^{\mu x}}{1 + c \cdot e^{\mu x}},$$

missä b , c ja μ ovat aineistosta estimoitavat vakiot. (Doray, 2008.)

Suomessa yleisesti käytetyt kuolevuusmallit, Makeham ja Gompertz, on havaittu käyttökelpoisiksi kuvaamaan kuolevuutta ja ne ovat edelleenkin käytössä kuolevuuden mallintamisessa. Näiden mallien suurimpina etuina voidaan pitää yksinkertaisuutta sekä melko hyvää tarkkuutta kuolevuuden kuvaamisessa. (Mäkinen, 2004.) Tutkielmassa naisten kohorttikohtaisiin henkivakuutuskuolevuuksiin sovitaan funktio, joka syntymävuoden mukaan määrää kuolevuuden ryhmäeläkevakuutuksen osalta. Funkti Sovite muodostetaan edellä esitettyä Makehamin mallia käyttäen.

Jos henkilö on x vuotta vanha, hänen ikänsä on välillä $[x, x + 1)$. Samoin jos tapahtuma sattuu kalenterivuonna t , tapahtuu se välillä $[t, t + 1)$. Usein on tiedossa kuolemien lukumäärä tietyssä ajanjaksona ja vastaavasti elossa olevien henkilöiden lukumäärä vastaavana ajanjaksona kohorteittain. Johdetaan kuolemistodennäköisyydet q_x eli vuoden alussa x -ikäisen henkilön todennäköisyys kuolla vuoden aikana. Kohorttikohtaista kuolevuutta kuvaava m_x -luku on määritelty kaavassa (4.1). Olkoon a_x keskiarvo ajasta, jonka $[x, x + 1)$ -ikäisinä kuolleet henkilöt ovat eläneet vuonna t . Asetetaan $a_x = 0.5$ jokaiselle kohortille. Tällöin voidaan laskea kuolemistodennäköisyydet q_x lukujen m_x ja a_x avulla kaavasta

$$(4.2) \quad q_x = \frac{m_x}{1 + (1 - a_x) \cdot m_x} = \frac{m_x}{1 + 0.5 \cdot m_x},$$

missä $x = 0, 1, 2, \dots, 100$. Olkoon käänteisesti p_x todennäköisyys pysyä elossa iästä x ikään $x + 1$, jolloin

$$p_x = 1 - q_x.$$

Olkoon alkukanta iässä 0 suuruudeltaan l_0 , jolloin elossa olevien henkilöiden lukumäärä iässä x on

$$l_x = l_0 \cdot \prod_{i=0}^{x-1} p_i.$$

(Kolehmainen, 2009.)

4.2 Korkomallit

Henkivakuutusyhtiöille korkojen aiheuttamat riskit ovat usein merkittäviä. Niitä mitataan esimerkiksi ennakoivan vakavaraisuuden tai Solvenssi II:n mukaisissa vakavaraisuuskehikoissa. Jos yhtiö kuitenkin kaipaa tarkempaa käsitystä korkoriskeistä, esimerkiksi omaa riskienhallintaansa tai näiden vakavaraisuuskehikoiden mukaisia sisäisiä malleja varten, on yhtiön perusteltua analysoida sekä koron vaihtelua että sen mukanaan tuomia riskejä itse. Tällöin yhtiö voi hyödyntää esimerkiksi stokastisten aikarakenneyhtälöiden tai differentiaaliyhtälöiden avulla saatuja korkomalleja. (Saraste, 2009.)

Tuottokäyrä kuvaa koron ja maturiteettien välisen riippuvuuden graafisesti. Tuottokäyrät kehittyvät dynaamisesti, joten niillä ei ole ainoastaan poikkileikkauksellinen vaan myös ajallinen ulottuvuus. (Diebold & Rudebusch, 2013.) Kaksi käytetyintä lähestymistapaa aikavälin rakenteen mallintamiseen ovat arbitraasittomuuteen (no-arbitrage) perustuvat mallit sekä tasapainomallit (equilibrium). Arbitraasimahdollisuus tarkoittaa, että saadaan voittoa sopivilla operaatioilla positiivisella todennäköisyydellä, mutta todennäköisyys jäädä tappiolle on nolla (Nyrhinen, 2012). Arbitraasittomuuteen perustuvia malleja yhdistää pyrkimys mallintaa aikavälin rakennetta täydellisesti tietyssä ajanhetkenä. Tasapainomallit taas keskittyvät mallintamaan hetkellisen koron dynamiikkaa. Tässä tutkielmassa ei kuitenkaan käytetä kumpaakaan lähestymistapaa vaan valitaan Nelson-Siegel (1987) -mallista Dieboldin ja Lin kehittämä dynaaminen versio, joka tiivistää koko tuottokäyrän kolmiulotteiseksi parametrivektoriksi, joka muuttuu dynaamisesti. Nämä parametrit voidaan tulkita faktoreiksi, joille ehdotetaan ja estimoidaan autoregressiivisiä malleja ja sitten faktoreita ennustamalla ennustetaan tuottokäyrää. Mallin vahvuutena voidaan pitää sen yksinkertaisuutta sekä mahdollisuutta hyvinkin tarkkaan estimointiin. (Diebold & Li, 2006.)

4.2.1 Dynaaminen Nelson-Siegel-korkomalli

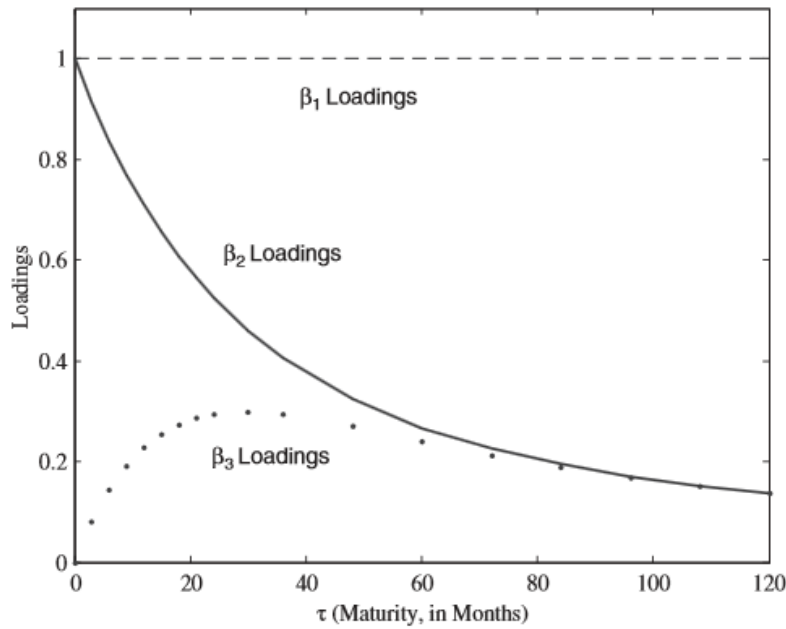
Dynaamisen Nelson-Siegel -mallin (DNS) lähtökohtana on sovittaa epälineaarinen käyrä tuottokäyrään jokaisella ajanhetkellä t . Hetkellisen Forward-koron eli termiinikoron käyräksi oletetaan

$$f_t(\tau) = \beta_{1t} + \beta_{2t}e^{-\lambda_t\tau} + \beta_{3t}\lambda_t\tau e^{-\lambda_t\tau},$$

missä τ on tasoa $f_t(\tau)$ vastaava maturiteetti. Termiinikorko määrittää etukäteen sen koron, joka tullaan myöhemmin maksamaan lainattavalle pääomalle (Korhonen, 2009). Tuottokäyrä voidaan rakentaa havaituilla maturiteeteilla estimoitujen termiinikorkojen avulla. Käytetään korkomallin pohjana Nelson-Siegel parametrisaatiota, joka tuottaa annetuilla parametreilla koko tuottokäyrän. Vastaava tuottokäyrän funktio on

$$(4.3) \quad y_t(\tau) = \beta_{1t} + \beta_{2t} \left(\frac{1 - e^{-\lambda_t\tau}}{\lambda_t\tau} \right) + \beta_{3t} \left(\frac{1 - e^{-\lambda_t\tau}}{\lambda_t\tau} - e^{-\lambda_t\tau} \right).$$

Malli ottaa huomioon tuottokäyrän muotoon keskeisimpänä vaikuttavat tekijät, joita ovat tuottokäyrän taso β_{1t} (level), kaltevuus β_{2t} (slope) sekä kaarevuus β_{3t} (curvature). Kaltevuutta voidaan pitää ”lyhytaikaisena faktorina”, joka vaikuttaa enimmäkseen lyhytaikaisiin tuottoihin ja kaarevuutta ”keskipitkän aikavälin faktorina” vaikuttaen pääosin keskipitkiin tuottoihin. Toisin kuin kaksi muuta faktoria, tuottokäyrän taso vaikuttaa pitkäaikaisiin tuottoihin ja sitä voidaankin kutsua ”pitkäaikaiseksi faktoriksi”. (Diebold & Rudebusch, 2013.)



Kuvio 5. Tuottokäyrän $y_t(\tau)$ faktorilataukset β_1 , β_2 ja β_3 , kun $\lambda_t = 0.0609$. Vaaka-akselilla maturiteetit τ on esitetty kuukausissa ja pystyakselilla annetaan faktoreiden latauskertoimet. (Diebold & Li, 2006)

Kuten kuviosta 5 nähdään, tuottokäyrän taso β_{1t} on vakio arvoltaan yksi, eikä se laske nolnaan missään vaiheessa. Tuottokäyrän kaltevuus β_{2t} on funktio, joka saa alkuarvoksi yksi, mutta laskee monotonisesti ja nopeasti nolaa kohti ja tuottokäyrän kaarevuus β_{3t} on funktio, joka taas alkaa nollasta, kasvaa mutta laskee takaisin nolaa kohti. Mallissa on vielä neljäs parametri λ_t , joka säätelee eksponentiaalista vaimenemisnopeutta; pienet λ_t :n arvot saavat aikaan hidasta vaimenemista ja sovittavat käyrän usein paremmin pitkillä maturiteeteilla kun taas suuret λ_t :n arvot aiheuttavat nopeaa vaimenemista ja sovittavat käyrän paremmin lyhyillä maturiteeteilla. Parametri λ_t säätelee myös, missä kohtaa parametrin β_{3t} lataus saavuttaa maksiminsa. Pitkäaikainen faktori β_{1t} määrää tuottokäyrän tason ja voidaan helposti osoittaa, että $y_t(\infty) = \beta_{1t}$. Näin ollen kasvu parametrissa β_{1t} nostaa kaikkia tuottoja sillä lataus on identtinen kaikissa maturiteeteissa ja muuttaa siksi tuottokäyrän tasoa. Lyhytaikainen faktori β_{2t} voidaan määritellä tuottona $y_t(120) - y_t(3)$, eli kymmenen vuoden tuottona vähennettynä kolmen kuukauden tuotolla. Keskipitkän aikavälin faktori β_{3t} voidaan määritellä $2y_t(24) - y_t(3) - y_t(120)$, eli kaksinkertaisena kahden vuoden tuottona vähennettynä kolmen kuukauden ja kymmenen vuoden tuotoilla. (Diebold & Li, 2006.)

Makroekonomisista muuttujista inflaatio on liitettävissä tuottokäyrän tasoon kun taas kaltevuus riippuu taloussuhdanteesta. Kaarevuus taas ei vaikuttaisi olevan suoranaisesti liitettävissä yhteenkään makroekonomiseen muuttujaan. (Diebold & Rudebusch, 2013.)

Muutetaan vielä mallin (4.3) merkintätapa painottamaan tulkintaa mallin faktoreista

$$y_t(\tau) = l_t + s_t \left(\frac{1 - e^{-\lambda_t \tau}}{\lambda_t \tau} \right) + c_t \left(\frac{1 - e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau} \right),$$

jossa $t = 1, \dots, T$ ja $\tau = 1, \dots, N$. Mallissa (4.3) l viittaa mallin tasoon (level), s kaltevuuteen (slope) ja c kaarevuuteen (curvature). Satunnaisten virhetermien lisääminen deterministiseen DNS-käyrään tuottaa yhtälön, joka satunnaisesti yhdistää joukon N tuottoja kolmeen ei-havaittuun tuottofaktoriin

$$(4.4) \quad y_t = \Lambda f_t + \varepsilon_t,$$

missä muuttujat ovat

$$y_t = \begin{pmatrix} y_t(\tau_1) \\ y_t(\tau_2) \\ \vdots \\ y_t(\tau_N) \end{pmatrix}, \quad f_t = \begin{pmatrix} l_t \\ s_t \\ c_t \end{pmatrix}, \quad \varepsilon_t = \begin{pmatrix} \varepsilon_t(\tau_1) \\ \varepsilon_t(\tau_2) \\ \vdots \\ \varepsilon_t(\tau_N) \end{pmatrix}$$

ja parametrien matriisi on muotoa

$$\Lambda = \begin{pmatrix} 1 & \frac{1 - e^{-\lambda_1 \tau}}{\lambda_1 \tau} & \frac{1 - e^{-\lambda_1 \tau}}{\lambda_1 \tau} - e^{-\lambda_1 \tau} \\ 1 & \frac{1 - e^{-\lambda_2 \tau}}{\lambda_2 \tau} & \frac{1 - e^{-\lambda_2 \tau}}{\lambda_2 \tau} - e^{-\lambda_2 \tau} \\ \vdots & \vdots & \vdots \\ 1 & \frac{1 - e^{-\lambda_N \tau}}{\lambda_N \tau} & \frac{1 - e^{-\lambda_N \tau}}{\lambda_N \tau} - e^{-\lambda_N \tau} \end{pmatrix},$$

missä $t = 1, \dots, T$. Jokaista tuottoa $y_t(\tau)$ määrittävät osaksi yleiset faktorit l_t , s_t sekä c_t ja osaksi faktori $\varepsilon_t(\tau)$ satunnaisvirheille. Määritellään seuraavaksi yleisten faktorien dynamiikkaa eli siirtymäfunktio. Käytetään ensimmäisen kertaluvun vektoriautoregressiota

$$(4.5) \quad (f_t - \mu) = A(f_{t-1} - \mu) + \eta_t,$$

missä muuttujat ovat

$$f_t = \begin{pmatrix} l_t \\ s_t \\ c_t \end{pmatrix}, \quad \eta_t = \begin{pmatrix} \eta_t^l \\ \eta_t^s \\ \eta_t^c \end{pmatrix}$$

ja parametrivektori ja -matriisi ovat

$$\mu = \begin{pmatrix} \mu^l \\ \mu^s \\ \mu^c \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Selvästi μ on faktorikeskiarvo ja A määrää faktoridynamiikat. Määritetään vielä mittaus- ja muutoshäiriöiden kovarianssirakenne. Tehdään perusoletus, että valkoisen kohinan mittaus- ja muutoshäiriöt ovat ortogonaalisia toisiinsa ja alkuperäiseen tasoon nähden:

$$(4.6) \quad \begin{pmatrix} \eta_t \\ \varepsilon_t \end{pmatrix} \sim WN \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} Q & 0 \\ 0 & H \end{pmatrix} \right),$$

$$(4.7) \quad \begin{aligned} E(f_0 \eta_t) &= 0, \\ E(f_0 \varepsilon_t) &= 0. \end{aligned}$$

Normaalisuutta ei vaadita, joten ennusteiden ja estimaattoreiden, jotka perustuvat oletuksiin (4.4)-(4.7), tehokkuus vaihtelee riippuen onko normaalius saavutettu vai ei. (Diebold & Rudebusch, 2013.)

Toimivan tuottokäyrän dynamiikan mallintamisen on vastattava paitsi aiemmin käsiteltyjä määritelmiä koskien tuottokäyrän keskimääräistä muotoa sekä erilaisia muotoja eri aikoina, mutta myös näiden muotojen dynaamista kehitystä. Katsotaan tarkemmin joitain tärkeimpiä määritelmiä ja sitä, kuinka dynaaminen malli kykenee toistamaan niitä:

1. *Keskimääräinen tuottokäyrä on kasvava ja konkaavi.*

Tutkimuksen puitteissa keskimääräinen tuottokäyrä on sellainen, joka vastaa parametrien β_{1t} , β_{2t} ja β_{3t} keskimääräisiä arvoja. Periaatteessa on täysin mahdollista, että tuottokäyrä voi olla kasvava ja konkaavi.

2. *Tuottokäyrä voi saada useita erilaisia muotoja, joita ovat esimerkiksi nouseva ja laskeva sekä käyrä, jolla on kyttyrä ("humped" yield curve) tai käänteinen kyttyrä.*

Tutkimuksen tuottokäyrä voi saada kaikki nämä muodot. Kuinka selkeästi ja kuinka usein tuottokäyrän muoto vaihtuu, riippuu muuttujien β_{1t} , β_{2t} ja β_{3t} vaihtelusta.

3. *Korkojen sopeutuminen shokkeihin on hitaampaa kuin korkomarginaalin.*

Korkomarginaali vastaa vahvaa pysyvyyttä β_{1t} kun taas korot vastaavat heikompaa pysyvyyttä β_{2t} .

4. *Lyhytaikaiset korot vaihtelevat herkemmin kuin pitkäaikaiset korot.*

Tutkimuksen puitteissa tällä viitataan faktorilatauksiin: lyhytaikainen tuottokäyrä riippuu positiivisesti muuttujista β_{1t} ja β_{2t} kun taas pitkäaikainen riippuu ainoastaan muuttujasta β_{1t} .

5. *Pitkät korot sopeutuvat shokkeihin hitaammin kuin lyhyet korot.*

Pitkät korot riippuvat ainoastaan muuttujasta β_{1t} . Jos β_{1t} on faktoreista pysyvin, pitkät korot tulevat sopeutumaan shokkeihin hitaammin kuin lyhyet korot. (Diebold & Li, 2006.)

Pitkien ja lyhyiden korkojen välisen eron, eli korkomarginaalin, käyttäytyminen on olennainen tekijä tuottokäyrän kaltevuuden määräytymisessä. Tuottokäyrän kaltevuus heijastaa markkinoiden odotuksia ja onkin tärkeä mittari kuvaamaan erilaisia riskejä. (Diebold & Li, 2006.)

4.2.2 Nelson-Siegel-korkomalli

Nelson ja Siegel kehittivät vuonna 1987 mallin, jonka jako tekijöihin eroaa hieman dynaamisesta mallista. Alkuperäisen Nelson-Siegel-mallin tasokäyrä voidaan määritellä seuraavasti

$$y_t(\tau) = b_{1t} + b_{2t} \left(\frac{1 - e^{-\lambda_t \tau}}{\lambda_t \tau} \right) - b_{3t} (e^{-\lambda_t \tau}),$$

missä $b_{1t} = \beta_{1t}$, $b_{2t} = \beta_{2t} + \beta_{3t}$ ja $b_{3t} = \beta_{3t}$. Dynaamista mallia pidetään kuitenkin parempana, sillä muun muassa latauksilla $(1 - e^{-\lambda_t \tau})/\lambda_t \tau$ ja $e^{-\lambda_t \tau}$ on samankaltainen monotonisesti laskeva muoto, mikä vaikeuttaa faktoreille b_2 ja b_3 annettavia erilaisia tulkintoja. Dynaamisessa mallissa lataukset ovat selkeästi erimuotoiset ja faktoreilla on vastaavasti erilaiset tulkinnat, jotka ilmenevät tasona, kaltevuutena ja kaarevuutena. (Diebold & Rudebusch, 2013.)

4.3 Solvenssi II

Euroopan Unionissa on meneillään laaja uudistushanke, joka koskee vahinko- ja henkivakuutusyritysten vakavaraisuusvalvontaa ja -säätelystä. Tavoitteena on luoda harmonisoitu riskiperusteinen sekä kokonaisvaltainen vakavaraisuuskehikko, joka edistää EU:n sisäistä kilpailua sekä pääomien tehokasta hyödyntämistä mutta myös yritysten omaa riskienhallintaa ja tätä kautta tehostaa vakuutettujen etujen turvaamista. Tämä niin kutsuttu Solvenssi II -direktiivi korvaa nykyiset säännöt EU:ssa vuonna 2014 ja sen on tarkoitus uudistaa vakuutusyritysten vakavaraisuutta koskevia säännöksiä sekä yhdistää useita vakuutustoimintaa koskevia direktiivejä yhdeksi tiiviiksi kokonaisuudeksi. Suomessa kyseiset muutokset koskevat ensisijaisesti vahinko-, henki- ja jälleenvakuutusyhtiöitä. (Finanssivalvonta, 2014.) Tässä osiossa käydään lyhyesti läpi muutamia Solvenssi II:n uusia keskeisiä ominaisuuksia, sillä osa tutkielmaan sisältyvistä laskelmista perustuu teorialtaan uudistuksen säännöksiin.

4.3.1 SCR -pääomavaatimus

Solvenssi II -direktiivin vakavaraisuussäännökset perustuvat markkina-arvoiseen taseeseen, jossa oma varallisuus *OF* sisältää omaisuuden ja velkojen erotuksen. Sitä kuinka *OF* tai nettovarallisuus muuttuu erilaisissa riskiskenaarioissa, analysoidaan kullekin riskityypille erikseen. Vakavaraisuuslaskennan lähtökohtana toimii SCR-pääomavaatimus, jonka tarkoitus on asettaa yrityksen oma varallisuus vähintään sellaiselle tasolle, jolla vakuutuksenottajien etujen voidaan sanoa olevan kohtuullisella tasolla turvattu. Jotta pääomavaatimukset täyttyisivät, on vakuutusyrityksellä jatkuvasti oltava riittävästi hyväksyttävää omaa varallisuutta. (Tiainen, 2011.) SCR-pääomavaatimuksella pyritään määrittelemään pääomamäärä, joka vakuutusyrityksellä on oltava ennakoimattomien tappioiden kestämiseksi. (Ronkainen, 2012.)

SCR-pääomavaatimus voidaan laskea kahdella eri tavalla, joko standardikaavalla tai yrityksen omalla sisäisellä mallilla. Näistä ensimmäisen tarkoituksena on kuvata useimpien vakuutusyritysten riskiprofiilia. Joissain tapauksissa tämä lähestymistapa ei kuitenkaan anna tarpeeksi oikeaa kuvaa, jolloin yritys voi laskea SCR-pääomavaatimuksensa käyttäen sisäistä mallia. (Tiainen, 2011.)

SCR-pääomavaatimus perustuu Value at Risk eli VaR-lukuun ja tarkoituksena on rajata todennäköisyys vararikolle Solvenssi II:n mukaiseen 0,5 prosenttiin 12 kuukauden ajanjaksolla. Pääomavaatimuksessa lasketaan siis 99,5 prosentin kvantiili eli niin kutsuttu VaR-luku vakuutusyhtiön oman varallisuus-

den muutokselle. Toisin sanoen SCR-pääomavaatimus on mitoitettu niin, että sen täytyessä vakuutusyrittäjä kykenee kohtaamaan merkittäviäkin tappioita samalla, kun se pystyy huolehtimaan vastuistaan ja velvoitteistaan vakuutuksenottajia sekä edunsaajia kohtaan (Tiainen, 2011).

4.3.2 Value at Risk

Talousmarkkinoilla on useita erilaisia riskejä, joista liiketoimintariski, luottoriski sekä markkinariski ovat talousriskien kolme pääkategoriaa (Tsay, 2010). Maksujärjestelyissä ja -ohjeissa tapahtuvia riskejä kutsutaan liiketoimintariskeiksi. Luottoriskillä taas tarkoitetaan riskiä, joka syntyy vastapuolen kyvyttömyydestä suorittaa omia velvoitteitaan ja markkinariskillä tulevaisuuden tuottojen epävarmuutta, mikä johtuu markkinaolosuhteiden muutoksista. Vakuutusyhtiöiden näkökulmasta riski mitataan yleensä vuoden lopussa mahdollisena arvonalenemisena (Ufer, 1996). Vakuutus- ja rahoitusmarkkinoilla on käytössä useita riskienhallinnan välineitä, joista yleisimmin käytetty on VaR. Se on esimerkiksi pohjana Solvenssi II:n SCR-pääomavaatimuksessa mutta käytössä myös erityisesti vakuutusyhtiöiden ja pankkien riskienhallinnassa. (Ronkainen, 2012.) Menetelmä otettiin käyttöön 1980-luvun lopussa, mutta sen suosio kasvoi huomattavasti 1990-luvulla J.P. Morganin RiskMetrics-järjestelmän ansiosta. Se tiivistää riskin yhteen helposti vertailtavaan valuuttamääräiseen lukuun ja näin ollen mahdollistaa eri yksiköiden tai toimintojen riskiarvion vertailun. On kuitenkin huomioitava, ettei VaR-luku kerro mitään tappion suuruudesta vaan sen tarkoitus on vain kertoa, kuinka todennäköisesti VaR-luku ylitetään. VaR-luku on liitettävissä pääosin markkinariskiin, mutta käsite soveltuu myös muun tyyppisiin riskeihin.

VaR-lähestymistavan ensimmäinen askel on päättää luottamustaso $1 - \alpha$ sekä aikaväli t , joka kannattaa valita samaksi kuin tarkastelun kohteena olevalla muuttujalla. Jotta VaR-luvun ylittävät tapaukset jäävät mahdollisimman vähäisiksi, on luottamustaso tarkoituksenmukaista valita mahdollisimman suureksi. On kuitenkin huomioitava, että mitä suurempi luottamustaso valitaan, sitä suuremmaksi riskienusteen epätarkkuus kasvaa, sillä samalla havaintojen määrä vähenee. Tämän vuoksi erittäin suurten luottamustasojen omaavia VaR-lukuja on tulkittava varovaisesti. (Jorion, 1997.) Aikaväli taas voi vaihdella tunneista vuosiin tarkoituksen mukaan, mutta yleensä ajanjaksona käytetään yhtä päivää tai kuukautta.

VaR-luvun käytön perusidea on luoda vararahasto, joka on suurella todennäköisyydellä riittävä kompensoimaan mahdollisia haitallisia hintojen muutoksia. Menetelmän keskeinen ominaisuus on, että sen

avulla saadaan selville yläraja kokonaistappion suuruudelle, joka ei ylitä tietyllä todennäköisyydellä $1 - \alpha$. Salkun arvon W_t markkinariskin mittaaminen koostuu varojen tuotto-prosessien mallintamisesta sekä niiden riippuvuusrakenteesta aikana $(t, t + s]$. Saadusta tuottojen ja tappioiden todennäköisyysjakaumasta valitaan α -kvantiili miinusmerkkisenä, $\alpha \in (0, 1)$. Toisin sanoen vaatimuksena on, että

$$P(W_{t+s} - W_t + VaR_\alpha < 0) = \alpha.$$

VaR-luvun määrittämiseen on olemassa kolme päämenetelmää: delta-normaali, historialliseen simulointiin sekä Monte Carlo -simulointiin perustuva menetelmä. Näistä ensimmäinen on analyyttinen ja kaikista suoraviivaisin menetelmä VaR-luvun laskemiseen, mutta siinä perusoletuksena on aineiston normaalijakautuneisuus. Kaksi muuta ovat simulaatiotekniikoita. Historiallinen simulaatio käyttää aineistoa sellaisenaan ja tarkasteltava jakauma muodostuu aineistossa tapahtuneista aidosta muutoksista. Monte Carlo -simulaatio on näistä kolmesta monimutkaisin menetelmä, mutta toisaalta se mahdollistaa myös monimutkaisista optioista muodostuvan salkun tarkastelun. Kaikilla kolmella lähestymistavalla on omat vahvuutensa ja heikkoutensa, mutta niitä ei tule tarkastella kilpailevina menetelminä vaan vaihtoehtoina erilaisille tilanteille. (Malz, 2011.)

Myös muita kuin markkinariskejä voidaan lähestyä VaR-näkökulmasta, kuten Solvenssi II:n SCR-pääomavaatimuksen tapauksessa on tehty. VaR-prosessissa kuitenkin tyypillisesti käytetään monia yksinkertaistavia oletuksia, jotka saattavat merkittävästi vaikuttaa riskin mittaamisen tarkkuuteen. VaR-lähestymistavassa on merkittäviä rakenteellisia rajoituksia, kuten esimerkiksi delta-normaali-menetelmän normaalijakamaoletus. Lisäksi se mahdollistaa riippuvuustarkasteluissa ainoastaan lineaarisen riippuvuuden. Kuitenkin yhteisten todennäköisyysjakaumien kehittyneempien mallinnusten ansiosta nämä rajoitukset voidaan kiertää. VaR-lukua mitattaessa on myös olemassa muita pienempiä ongelmia, joten muita riskien mittaamisvälineitä on kehitetty. (Ronkainen, 2012.)

4.4 Elinikäinen eläke

Eläkevakuutusjärjestelmän yleinen ajatus on siirtää vakuutusmaksut ja sijoitustuotot eläkkeiksi. Prosessi on luonteeltaan stokastinen, sillä vain pieni osa järjestelmän osatekijöistä voidaan varmuudella tietää etukäteen. Prosessissa on siis epävarmuutta, mikä edellyttää tarvittavia riskienhallintamenetelmiä. Yleisesti ottaen riskienhallinta sisältää riskien tunnistamisen, analysoinnin sekä arvioinnin. Riskienhallinnan avulla löydetään myös parhaat menetelmät käsitellä näitä osa-alueita. Toimenpiteet voivat sisältää

riskien poistamista, jakamista, lieventämistä tai puskurointia. Eläkevakuutussektorilla kaikkia riskejä ei yleensä voida poistaa eläkejärjestelmästä mutta niitä voidaan jakaa tai lieventää monin eri tavoin sidosryhmien keskuudessa. Eläkevakuutuksessa voidaan nähdä olevan kolme eri sidosryhmää: rahoittaja, edunsaaja sekä vakuutuksenantaja. Rahoittajan osa on maksaa eläkevakuutus sopimuksen mukainen vakuutusmaksu. Vakuutuksenantajan vastuulla on tarjota sopimuksen mukaiset eläkevakuutuspalvelut, jotka sisältävät vakuutusmaksujen asettamisen riittäviksi ja varojen hallinnoimisen viisaasti, jotta eläkkeet voidaan maksaa sovitusti. Eläkkeelle siirtymisen jälkeen edunsaaja saa eläke-etuja, kun säästöt muutetaan eläkkeeksi. Edunsaaja voi myös olla rahoittaja tietyissä eläkejärjestelmissä. Jos vakuutus on otettu myös kuoleman varalta, voi edunsaajia olla useampia. (Ronkainen, 2012.)

Elinikäisen eläkkeen osalta joko vakuutuksenantaja tai rahoittaja tai molemmat kohtaavat pitkäikäisyysriskin. Vakuutusyhtiöt ja eläkerahastot voivat lieventää pitkäikäisyysriskiä esimerkiksi erilaisten jälleenvakuutusjärjestelyjen kautta. Vakuutusyhtiöt, jotka tarjoavat sekä henkivakuutuksia että elinkorjoja hyötyvät tästä, sillä tällöin kuolevuus- sekä pitkäikäisyysriski osittain kompensoivat toisiaan. (Ronkainen, 2012.)

Eläkettä maksetaan korkeintaan niin kauan kuin eläkettä nauttiva henkilö on elossa. Näin ollen käytettävän koron lisäksi eläkkeen pääoma-arvo riippuu myös kuolevuudesta. Mitä matalampi kuolevuus on, sitä enemmän eläkkeitä tulee maksettavaksi. (Kolehmainen, 2009.) Pääoma-arvokertoimia käytetään muun muassa vastuuvelan sekä vakuutusmaksujen laskennassa (Aho & Sankala, 2011). Vastuuvellalla tarkoitetaan tulevien korvausten pääoma-arvoa vähennettynä tulevien vakuutusmaksujen pääoma-arvolla (Nyrhinen, 2012). Vastuuvelan riittävyyden takaamiseksi eläkevakuutuksen osalta on kuolevuusoletuksen oltava riittävän alhainen, jotta vastuuta tulee varattua tarpeeksi eläkettä nauttivien vakuutettujen tai edunsaajien maksamista varten. (Kolehmainen, 2009.)

Oletetaan, että eläkeläisten kohortti on täsmälleen iältään $x = 65$ perusvuoden $t = 0$ alussa. Valitaan vapaasti kohortin koko N_0 , joka voi vaihdella esimerkiksi yhdestä tuhanteen. Kuolevuusmallina käytetään Makehamin mallia ikään $\omega = 100$ asti. Oletetaan, että elämä päättyy viimeistään tässä iässä. Maksetaan kohortille, joka koostuu henkilöistä $i \in \{1, \dots, N_0\}$ vuotuista yksikköeläkettä eli vuosittain kuolemaan asti maksettavaa yhden yksikön suuruista eläkettä vuosille $t = 0, 1, 2, \dots, 34$. Kuolevuutta simuloidaan niin, että oletetaan kuolleiden lukumäärän noudattavan binomijakaumaa $\text{Bin}(N_t, q_{65+t})$, missä N_t on eläkeläisten lukumäärä vuoden t alussa ja q_{65+t} kuoleman todennäköisyys iässä $65 + t$ kaavan (4.2) mukaisesti. Eläkkeen satunnainen pääoma-arvo saadaan kaavasta

$$(4.8) \quad PV = \sum_{\tau=1}^{35} v_0(\tau) N_{\tau-1},$$

missä $v_0(\tau)$ on kaavassa (2.1) määritelty diskonttaustekijä. PV on satunnaismuuttuja, jonka satunnaisuus johtuu henkiinjäämisen epävarmuudesta sekä stokastisista koroista. Pääoma-arvo, jota yleensä käytetään teoreettisena nettovakuutusmaksuna, saadaan laskemalla odotusarvo $E(PV)$. Lisäksi tutkielmassa lasketaan satunnaisten pääoma-arvojen jakaumasta 99,5 % kvantiili eli VaR.

Solvenssi II:ssa vastuuvelan ja niitä kattavien korkosijoitusten maturiteettien ero näkyy korkeampana vakavaraisuusvaatimuksena. Tämä johtuu taseen liikkumisesta markkinakorkokäyrän mukaisesti. Korkokäyrän muutos aiheuttaa vakuutusyhtiölle voittoa tai tappiota, jos taseen eri puolilla kassavirrat ovat erilaisia. Tappiota varten vaaditaan riittäviä pääomia, ja pääomavaatimus mitoitetaan varojen ja velkojen erotuksena eli oman pääoman OF VaR-lukuna vuoden päähän arvioituna 99.5 % tasolla.

Yksinkertaisuuden vuoksi oletetaan, että varat sijoitetaan nollakuponkilainaan maturiteetilla τ . Tällöin vakuutusyhtiön oman varallisuus OF vuonna t on

$$(4.9) \quad OF(t) = \frac{[1 + y_{t-1}(\tau)]^\tau}{[1 + y_t(\tau - 1)]^{\tau-1}} PV_{t-1} - PV_t,$$

missä $[1 + y_{t-1}(\tau)]^\tau / [1 + y_t(\tau - 1)]^{\tau-1}$ kuvaa obligaatiosijoituksen arvon muutosta vuoden $t - 1$ aikana ja PV_t eläkkeen satunnaista pääoma-arvoa vuonna t . Tutkielmassa korkoja ennustetaan vuoden päähän, joten asetetaan vakuutusyhtiön oman varallisuuden muutos valitulla maturiteetilla tähän ajankohtaan. Aloitusajanhetkenä $t = 0$ oma varallisuus on $OF(0) = 0$, sillä ensimmäisenä vuonna kerätyt varat sijoitetaan mutta ne eivät vielä ehdi kerryttää tuottoa. Lisäksi ensimmäisenä vuonna vakuutusyhtiölle tulee maksettavaksi eläkkeitä koko populaatiolle, jolta varat on kerätty. Aloitusajanhetkenä eläkkeen satunnainen pääoma-arvo on

$$PV_0 = \sum_{\tau=1}^{35} v_0(\tau) N_{\tau-1},$$

ja vuoden päähän taas

$$PV_1 = N_0 + \sum_{\tau=1}^{34} v_1(\tau)N_{\tau}.$$

Tutkielman esimerkkilaskelmissa riskiä käsitellään laskemalla vakuutusyhtiön oman varallisuuden muutokselle vuoden aikana ($OF(t) - OF(t - 1)$) 0,5 % kvantiilin vastaluku eli VaR-luku. Laskuissa oletetaan lisäksi, että elossa olevan populaation eläkkeet maksetaan koko vuodelta, vaikka henkilö kuolisikin aiemmin samana vuonna.

5 AINEISTOJEN ANALYSOINTI

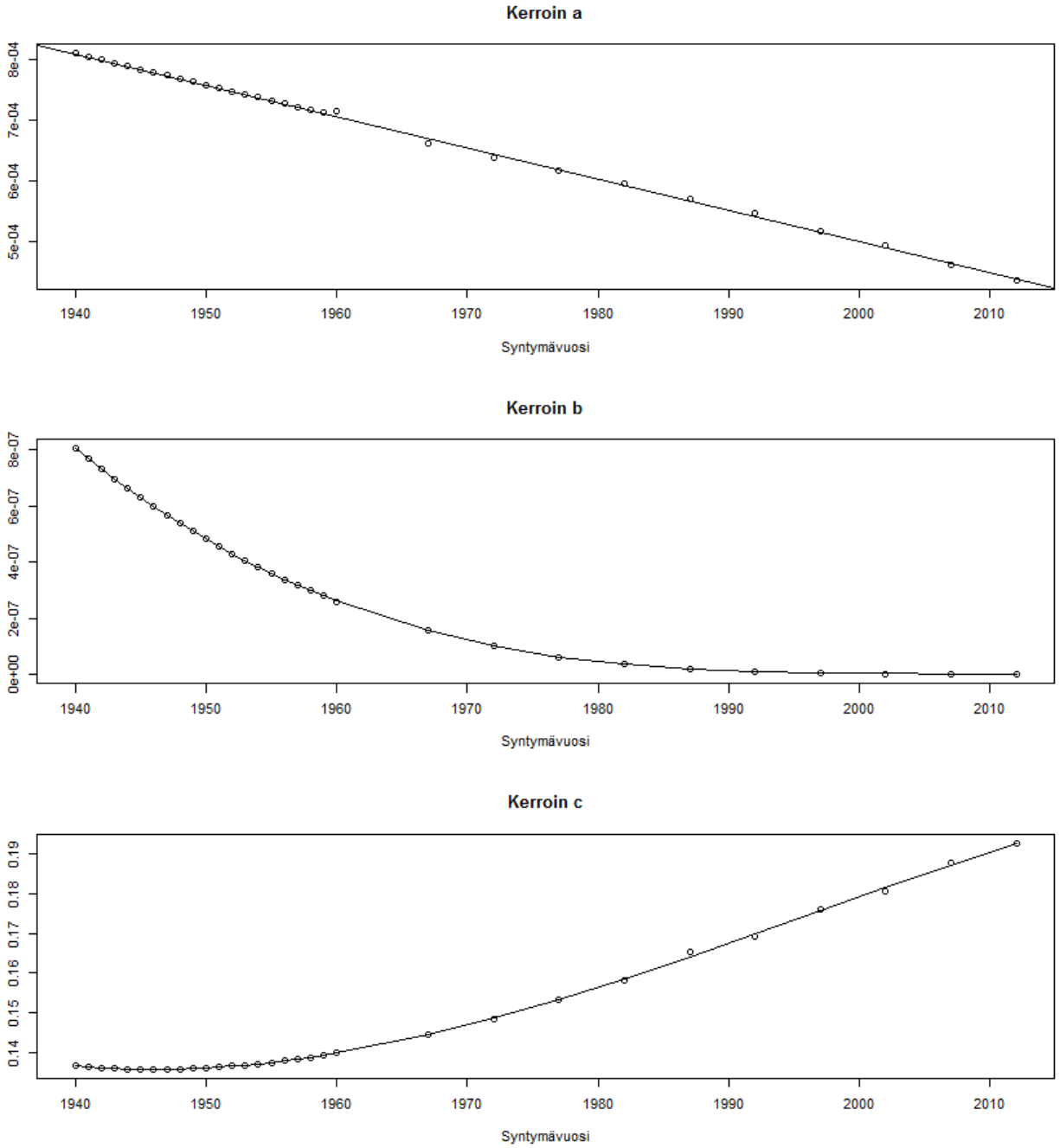
Tässä luvussa tarkastellaan kuolevuus- ja korkoaineistoja. Aluksi estimoidaan kuolevuusaineistoon sovitettava Makeham-kuolevuusmalli regressioanalyysin avulla. Tämän jälkeen siirrytään korkomalleihin sekä korkoaineiston mallintamiseen ja ennustamiseen aikasarja-analyysia soveltaen.

5.1 Makeham-kuolevuusmalli

Parhaan kuolevuusennusteen saa käyttämällä kohorttikohtaista kuolevuustaulua, mutta usein kuolevuus esitetään funktiomuodossa. Yksinkertaista Makeham-kuolevuusmallia ei kuitenkaan saada täysin tarkasti seuraamaan aineistoa, vaan sovituksen tekeminen lisää hieman mahdollista ennustevirhettä.

Sovitettaessa kuolevuusmallia on alle 40-vuotiaat poistettu aineistosta, sillä on havaittu, että tätä nuoremmissa ikäryhmissä ei ole suurta käytännön merkitystä ryhmäeläkevakuutusta tarkasteltaessa. Funktiot tehdään aineistoon, joka sisältää ikävuodet 40–100. Funktio sovitetaan kohorttiaineistoon ja tehdään valituille kohorteille vuodesta 1960 vuoteen 2012, minkä lisäksi malleja jatketaan ekstrapolomalla menneisyyteen aina vuoteen 1940. Tämän jälkeen funktiot kertoimille muodostetaan käyttäen koko käyrää vuodesta 1940 vuoteen 2012.

Kuviossa 6 on esitetty funktion kertoimien a , b ja c kuvaajat, mistä nähdään, että kaikkien kertoimien kuvaajat ovat monotonisia. Lisäksi kertoimen a kuvaaja voidaan olettaa lähes lineaariseksi ja sen muu-
tosta mallinnetaan ensimmäisen asteen lineaarisella regressiolla, selityssasteen ollessa 0,99919. Kertoimille b ja c etsitään parhaiten aineistoon sopiva funktio käyttämällä polynomiregressiota. Kolmannen asteen funktioiden havaittiin kuvaavan aineiston kertoimia riittävän tarkasti saaden selityssasteet 0,99999 ja 0,99960. Kaikille kolmelle kertoimelle muodostetaan funktiot käyttämällä argumenttina syntymävuotta. Selityssasteiden ja kuvion 6 avulla havaitaan myös, että sovitetut mallit istuvat hyvin mallinnettuun aineistoon. Mallinnuksessa käytetyt koodit esitetään liitteessä A sivulla 60.



Kuvio 6. Makehamin kuolevuusfunktion kertoimet a, b ja c syntymävuosittain.

Aiemmin luvussa 4.1 esitetty kohorttikohtainen Makeham-kuolevuusfunktio (4.1) voidaan yleistää muotoon:

$$(5.1) \quad \mu(x)_{sp} = a_{sp} + b_{sp} \cdot e^{c_{sp}x} = a_{sp} + e^{\ln(b_{sp}) + c_{sp}x},$$

missä x :llä tarkoitetaan ikää ja sp :llä sukupuolta [nainen]. Kun mallissa (5.1) $sp = nainen$, kertoimet saavat arvot

$$a_{nainen} = 0,010790991735695 - 5,14537051947357 \cdot 10^{-6} \cdot sv$$

$$b_{nainen} = 0,0256688686517275 - 3,84500600085219 \cdot 10^{-5} \cdot sv + 1,91985420209458 \cdot 10^{-8} \cdot sv^2 - 3,19536396250552 \cdot 10^{-12} \cdot sv^3$$

$$c_{nainen} = 1211,311076557 - 1,82192050833558 \cdot sv + 0,000912984605307542 \cdot sv^2 - 1,52403652766762 \cdot 10^{-7} \cdot sv^3,$$

missä sv = syntymävuosi.

5.2 DNS-mallin estimointi ja ennustaminen

Dynaamiselle Nelson-Siegel-korkomallille on olemassa useita estimointimenetelmiä, kuten esimerkiksi kaksivaiheinen menetelmä tai bayesiläinen data-analyysi (Diebold & Rudebusch, 2013). Tässä tutkielmassa käytetään kaksivaiheista menettelytapaa, jonka Diebold ja Li kehittivät vuonna 2006, ja sen mallinnuksessa käytetyt koodit esitetään liitteessä B sivulla 61.

Tarkastellaan kaksivaiheisen DNS:n tapausta, jossa λ_t on asetettu vakioksi kaikilla arvoilla t . Diebold ja Li asettivat aikafaktorin λ_t maksimoimaan keskipitkän aikavälin faktoria β_{3t} 30 kuukauden kohdalla, jolloin he onnistuivat muuntamaan ongelman epälineaarisesta sovittamisesta yksinkertaiseen lineaariseen regressioon. Asetetaan siis $\lambda_t = 0,0609$ vastaamaan Dieboldin ja Lin aikafaktoria. Menetelmän ensimmäisessä vaiheessa Nelson-Siegel-malli (4.3) sovitetaan jokaiselle ajanjaksolle $t = 1, \dots, T$ pienimmän neliösumman menetelmällä (OLS). Tämä vaihe tuottaa estimoitujen faktoreiden $\{\hat{l}_t, \hat{s}_t, \hat{c}_t\}_{t=1}^T$ kolmiulotteisen aikasarjan ja vastaavan N-ulotteisen sarjan virhetermeille $\{\hat{\varepsilon}_t(\tau_1), \hat{\varepsilon}_t(\tau_2), \hat{\varepsilon}_t(\tau_N)\}_{t=1}^T$. Ensimmäisen vaiheen toimivuus johtuu siitä, että dynaaminen malli tiivistää N-ulotteisen tuottoaikasarjan kolmiulotteiseksi tuottofaktoreiden $\{\hat{l}_t, \hat{s}_t, \hat{c}_t\}_{t=1}^T$ aikasarjaksi. Toisessa vaiheessa sovitetaan dynaaminen malli tuottofaktoreihin $\{\hat{l}_t, \hat{s}_t, \hat{c}_t\}_{t=1}^T$. Toinen vaihe tuottaa dynaamisten parametrien estimaatit määräten tuottofaktoreiden kehityssuunnan (siirtymäyhtälön parametrit) kuin myös virhetermien estimaatit. (Diebold & Rudebusch, 2013.)

Kaksivaiheisen estimoinnin hyödyt vakioidulla λ :lla ovat sen yksinkertaisuus, kätevyys ja numeerinen vakaus: mitään muuta ei edellytetä kuin lineaarista regressiota. Toki λ voidaan myös estimoida, jolloin ensimmäisen vaiheen pienimmän neliösumman menetelmän regressio muuttuu neliparametriseksi epälineaariseksi pienimmän neliösumman regressioiksi ja toisen vaiheen kolmiulotteinen dynaaminen malli tuottofaktoreille $\{\hat{l}_t, \hat{s}_t, \hat{c}_t\}_{t=1}^T$ muuttuu neliulotteiseksi dynaamiseksi malliksi faktoreille $\{\hat{l}_t, \hat{s}_t, \hat{c}_t, \hat{\lambda}_t\}_{t=1}^T$. (Diebold & Rudebusch, 2013.) Diebold ja Li kuitenkin asettivat λ_t :n maksimoimaan β_{3t} :n latausta 30 kuukauden kohdalla, mikä siis yksinkertaisti merkittävästi mallin sovittamista. Myös tässä tutkielmassa päädyttiin vakioimaan λ_t .

Taulukossa 1 esitellään tunnuslukuja aineistoon sovitetuista kuukausittaisista tuotoista sekä tuotto-käyrän kolmesta faktorista: tasosta, kaltevuudesta ja kaarevuudesta, jotta saadaan käsitys miten tuotto-käyrä käyttäytyy. Määritellään taso kymmenen vuoden tuotoksi, kaltevuus kymmenen vuoden ja kolmen kuukauden tuoton erotukseksi sekä kaarevuus kaksinkertaisen kahden vuoden tuoton ja kolmen kuukauden sekä kymmenen vuoden tuottojen erotukseksi.

Taulukko 1. Sovitetun tuottokäyrän tunnuslukuja

Maturiteetti	Keskiarvo	Mediaani	Keskihajonta	Minimi	Maksimi	Vinous	Huipukkuus
3	4,665	5,084	3,007	0,077	14,471	0,327	-0,120
6	4,782	5,074	3,084	0,020	14,582	0,327	-0,128
12	5,000	5,177	3,180	-0,051	14,722	0,342	-0,124
24	5,366	5,395	3,223	0,132	14,806	0,400	-0,083
36	5,647	5,595	3,174	0,375	14,781	0,461	-0,039
60	6,021	5,802	3,029	0,784	14,752	0,558	0,023
84	6,240	5,939	2,916	1,078	14,636	0,618	0,055
120 (taso)	6,425	6,022	2,815	1,347	14,501	0,666	0,074
Kaltevuus	1,760	1,821	1,165	-0,743	3,832	-0,170	-1,047
Kaarevuus	-0,358	-0,191	0,906	-2,413	2,739	-0,185	-0,228

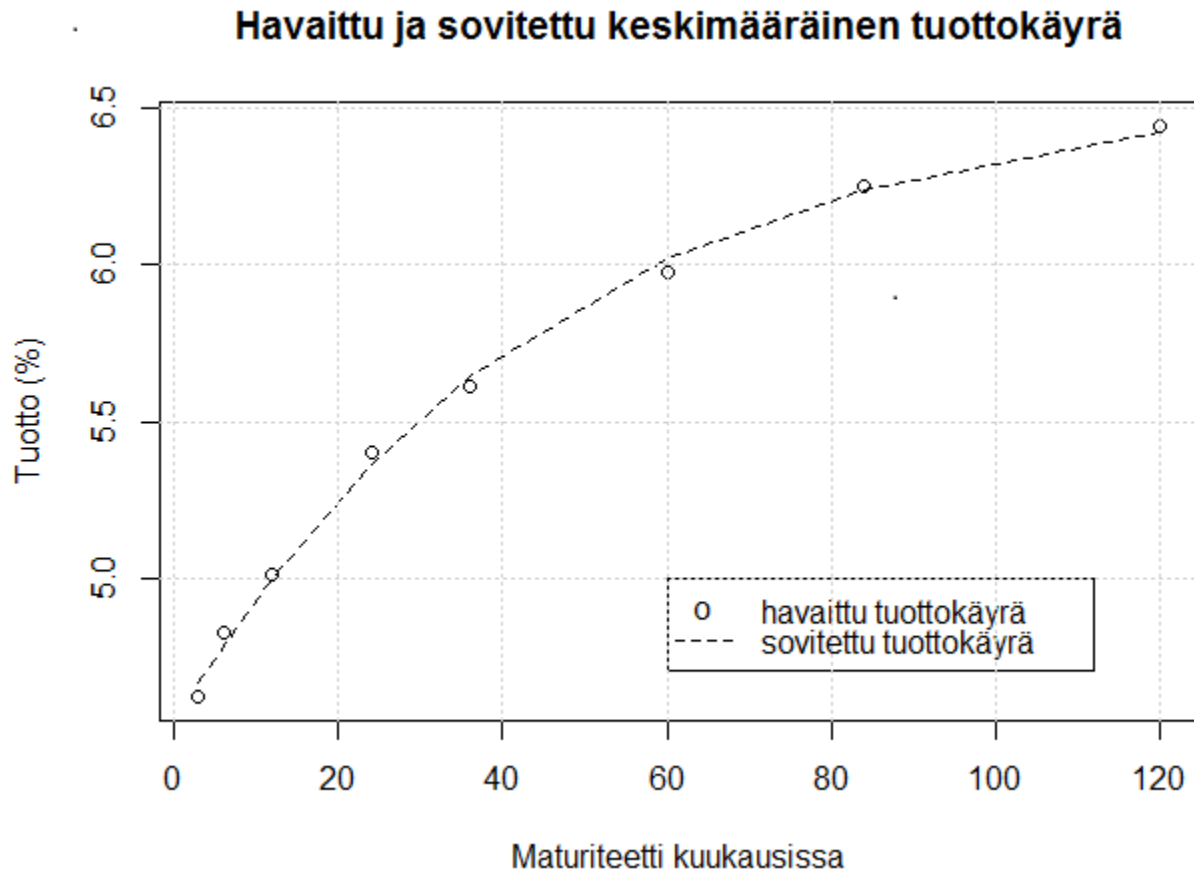
Taulukosta 1 nähdään, että maturiteettien tuottokorkojen jakaumat ovat hieman oikealle vinoja (*skewness*). Huipukkuudella (*kurtosis*) mitataan aineiston jakauman muotoa. Negatiivinen huipukkuus viittaa

tasaiseen jakaumaan ja positiivinen taas terävähuippuiseen jakaumaan. Jos aineisto olisi normaalijakautunut, huipukkuus saisi arvon nolla. Taulukosta 1 havaitaan, että kaikkien maturiteettien jakaumat ovat lähes normaalijakautuneita. Tuotot näyttäisivät kehittyvän maturiteettien kasvaessa samankaltaisesti, mikä näkyy myös vahvana riippuvuutena eri maturiteettien välillä. Riippuvuussuhteet on esitetty taulukossa 2.

Taulukko 2. Korrelaatiomatriisi eri maturiteettien tuottokoroille.

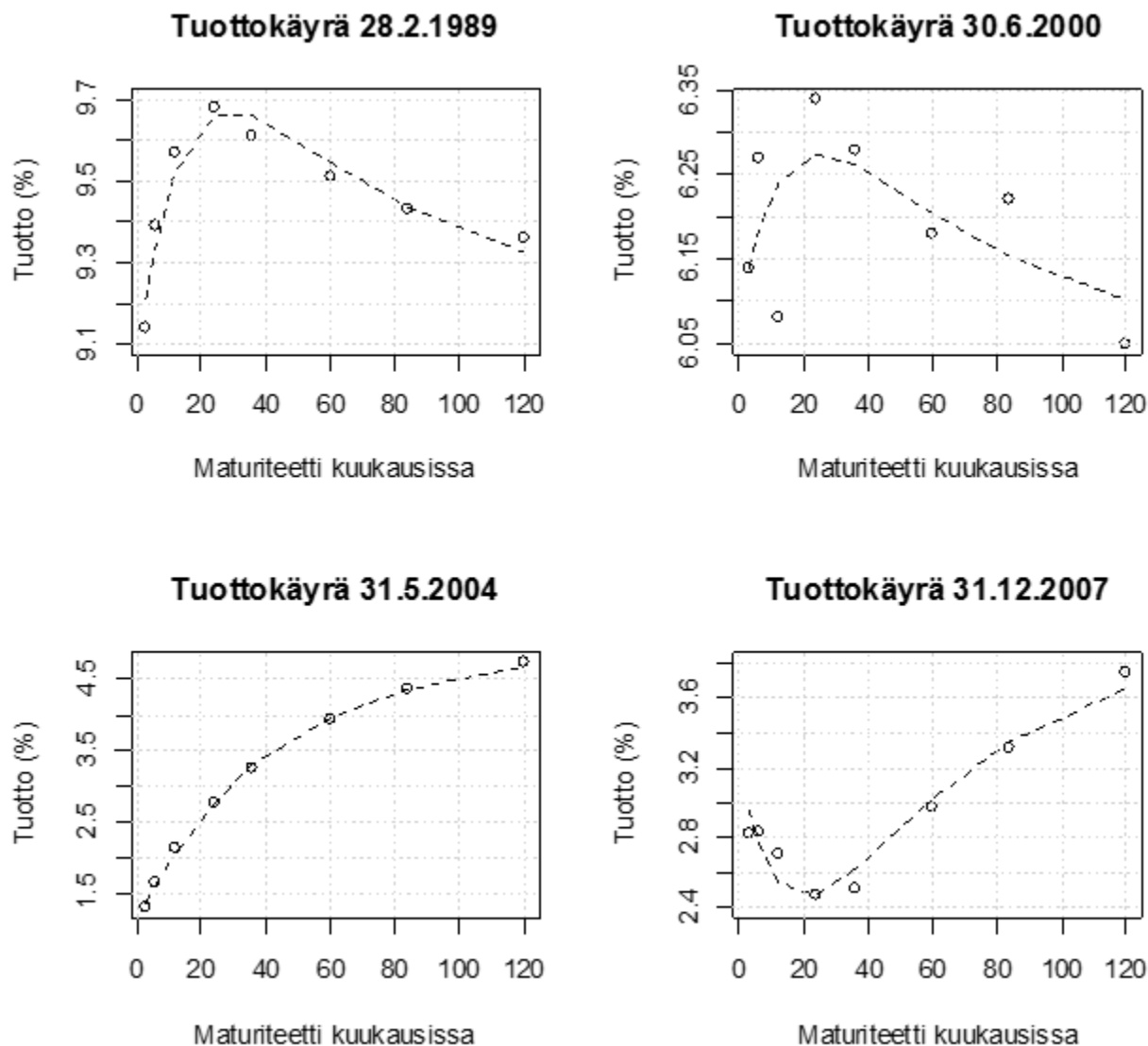
Maturiteetti	3	6	12	24	36	60	84	120
3	1							
6	0,9992	1						
12	0,9950	0,9982	1					
24	0,9842	0,9904	0,9968	1				
36	0,9737	0,9816	0,9909	0,9984	1			
60	0,9548	0,9645	0,9771	0,9901	0,9964	1		
84	0,9392	0,9498	0,9641	0,9806	0,9899	0,9983	1	
120	0,9220	0,9333	0,9491	0,9686	0,9807	0,9936	0,9985	1

Mallin sovittamista aineistoon voidaan tarkastella monelta eri kannalta. Ensinnäkin mallinnus voidaan toteuttaa vertaamalla havaittua ja sovitettua keskimääräistä tuottokäyrää keskenään, kuten kuviossa 7. Keskimääräiset tuottokäyrät on laskettu parametrien $\{\hat{\beta}_{1t}, \hat{\beta}_{2t}, \hat{\beta}_{3t}\}$ keskiarvoina eri maturiteetteina. Havaittu ja sovitettu tuottokäyrä näyttäisivät vastaavan toisiaan melko hyvin, joten malli vaikuttaisi tältä osin sopivalta.



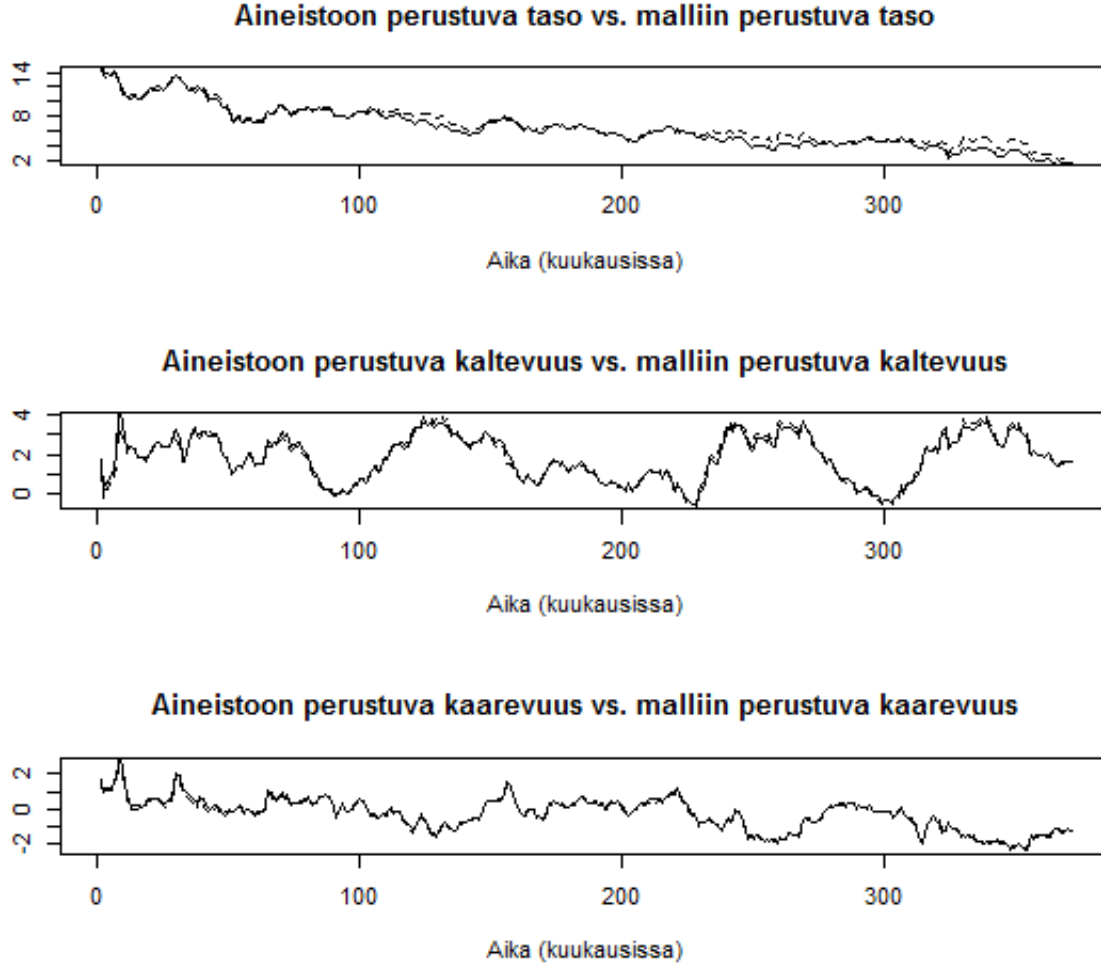
Kuvio 7. Havaittu (aineistoon perustuva) ja sovitettu (malliin perustuva) keskimääräinen tuottokäyrä.

Toiseksi mallinnus voidaan toteuttaa vertaamalla sovitettua kolmen faktorin tuottokäyrää eri ajankohtina havaittuihin tuottoihin kuten kuviossa 8. Kaikkien aineistossa olevien tuottokäyrien tarkastelu olisi kuitenkin melko työlästä, joten kuviossa 8 esitetään aineistosta löytynyttä neljää erimallista tuottokäyrää. Havaitut tuotot esitetään kuviossa pisteinä ja sovitetut tuottokäyrät katkoviivoina valituille neljälle kuukaudelle. Kuviosta 8 nähdään, että sovitetut tuottokäyrät kykenevät toistamaan monia tuottokäyrän muotoja, kuten ylös- sekä alaspäin kaltevuuden, kaarevuuden sekä käänteisesti kaarevuuden. Sovitetuilla käyrillä on kuitenkin vaikeuksia erityisesti sellaisten päivämäärien kohdalla, jolloin tuotot ovat hajanaisia sisältäen mahdollisesti useampia minimejä ja maksimeja.



Kuvio 8. Sovitetun mallin tuottokäyrät (katkoviiva) havaittujen tuottojen rinnalla valittuina kuukausina.

Kuviossa 9 esitetään aikasarjaestimaatit $\{\hat{\beta}_{1t}, \hat{\beta}_{2t}, \hat{\beta}_{3t}\}$ (katkoviiva) aiemmin määriteltyjen empiiristen muuttujien (kiinteä viiva): tason, kaltevuuden ja kaarevuuden rinnalla. Kuvio tukee väitettä, että mallin kolme faktoria todella vastaavat näitä muuttujia. Korrelaatiokertoimet estimoitujen faktorien ja empiiristen muuttujien kanssa ovat $\rho(\hat{\beta}_{1t}, l_t) \approx 0,988$, $\rho(\hat{\beta}_{2t}, s_t) \approx -0,989$ ja $\rho(\hat{\beta}_{3t}, c_t) \approx 0,998$, missä (l_t, s_t, c_t) ovat siis taso (level), kaltevuus (slope) ja kaarevuus (curvature).



Kuvio 9. Aineistoon perustuvat (kiinteä viiva) taso, kaltevuus ja kaarevuus vs. malliin perustuvat (katkoviiva) taso $\hat{\beta}_{1t}$, kaltevuus $-0.78\hat{\beta}_{2t}$ ja kaarevuus $0.37\hat{\beta}_{3t}$. Aineistoon perustuva taso määritellään kymmenen vuoden tuottona, kaltevuus kymmenen vuoden tuottona vähennettynä kolmen kuukauden tuotolla ja kaarevuus kaksinkertaisena kahden vuoden tuottona vähennettynä kolmen kuukauden ja kymmenen vuoden tuotoilla.

Taulukossa 3 annetaan kuvioita tukevia tunnuslukuja estimoiduille faktoreille $\hat{\beta}_{1t}$, $\hat{\beta}_{2t}$ ja $\hat{\beta}_{3t}$. Kuten nähdään myös kuviosta 10, ensimmäisen faktorin autokorrelaatiokerroin $\hat{\rho}(h)$ vähenee hitaammin ykkösestä kuin kahden muun faktorin, kun $|h|$ kasvaa. Kaikkien kolmen faktorin $\hat{\rho}(1)$ on kuitenkin hyvin lähellä ykköstä. Taulukon 3 arvot on laskettu koko aikasarjasta, ei siis vain 60 ensimmäisen kuukauden viiveistä, kuten kuviossa 10 on esitetty.

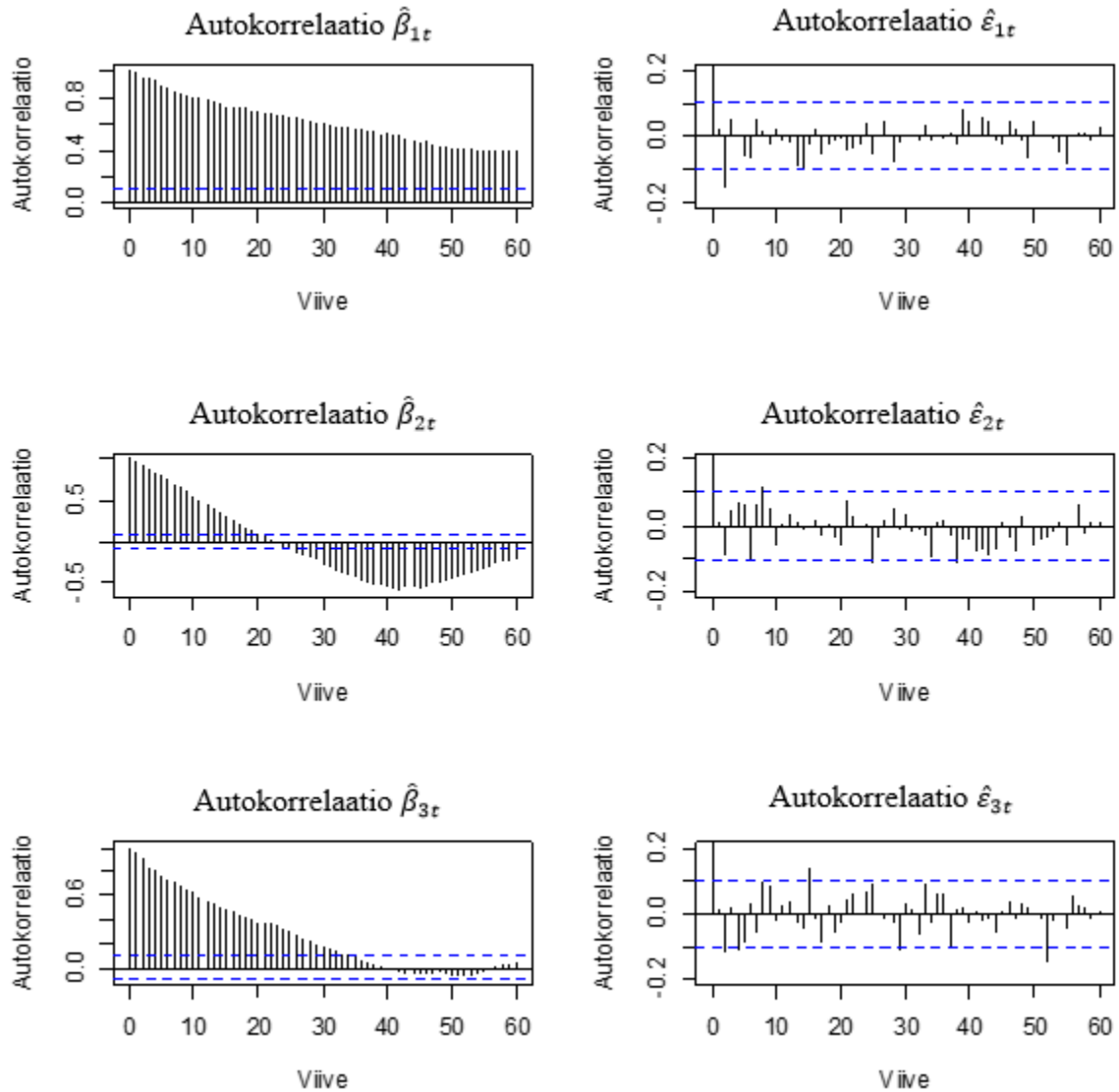
Soveltamalla pienimmän neliösumman menetelmää tuottoaineistoon saadaan aikasarjaestimaatit $\{\hat{\beta}_{1t}, \hat{\beta}_{2t}, \hat{\beta}_{3t}\}$ sekä niitä vastaavat jäännössarjat. Taulukko 3 ja kuvio 10 (vasen sarake) kuvaavat esti-

moituja faktoreita. Näiden kolmen faktorin autokorrelaatioista voidaan nähdä, että ensimmäinen faktori on kaikista pysyvin ja että kolmas faktori on pysyvämpi kuin toinen.

Taulukko 3. Sovitettuun Nelson-Siegel -tuottokäyrään perustuvia arvoja faktoreille $\hat{\beta}_{1t}$, $\hat{\beta}_{2t}$ ja $\hat{\beta}_{3t}$, kun $\lambda_t = 0.0609$. Kolme viimeistä saraketta sisältävät autokorrelaatiot 1, 12 ja 30 kuukauden viiveillä.

Faktori	Keskiarvo	Keskihajonta	Minimi	Maksimi	$\hat{\rho}(1)$	$\hat{\rho}(12)$	$\hat{\rho}(30)$
$\hat{\beta}_{1t}$	6,875	2,595	2,024	14,133	0,979	0,771	0,596
$\hat{\beta}_{2t}$	-2,333	1,570	-5,085	0,853	0,974	0,453	-0,260
$\hat{\beta}_{3t}$	-0,965	2,446	-6,512	7,400	0,957	0,560	0,171

Yleensä pitäisi pyrkiä valitsemaan sellainen malli, joka sisältää vähän parametreja ja on mahdollisimman yksinkertainen. Toisaalta taas mallilla tulisi olla riittävä yhteensopivuus aineiston kanssa. Parametrien määrän kasvaessa yhteensopivuus kasvaa, mutta toisaalta tällöin parametrien estimoinnin tarkkuus heikkenee. Kun parametrien estimaatit ovat heikkoja, käy ennustaminen epäluotettavaksi. Ennen kuin malli hyväksytään, tulisi tarkastella sen jäännössarjaa. Jäännössarja toimii estimaattina valkoiselle kohinalle, joten jos malli on hyvä, sillä pitäisi olla samat ominaisuudet kuin valkoisen kohinan sarjalla. Toisin sanoen mallin jäännössarjan on muistutettava valkoista kohinaa, jotta mallin voidaan olettaa olevan riittävä. Jäännössarjan valkoisuutta voidaan tarkastella erilaisten kuvaajien avulla, mutta on myös tarpeellista suorittaa testejä valkoisuuden varmistamiseksi. Testaamiseen voidaan käyttää useita samoja tapoja kuin itse aikasarjan testaamiseen, kuten Box-Ljung-testiä. Kuviossa 10 (oikea sarake) esitetään AR(2)-mallin sovite faktoreiden jäännössarjoille $\{\hat{\varepsilon}_{1t}, \hat{\varepsilon}_{2t}, \hat{\varepsilon}_{3t}\}$, mistä nähdään, että malli sopii hyvin estimoiduille faktoreille. Autokorrelaatiot ovat hyvin pieniä osoittaen, että mallit kuvaavat tarkalleen tason, kaltevuuden ja kaarevuuden ajallista riippuvuutta.



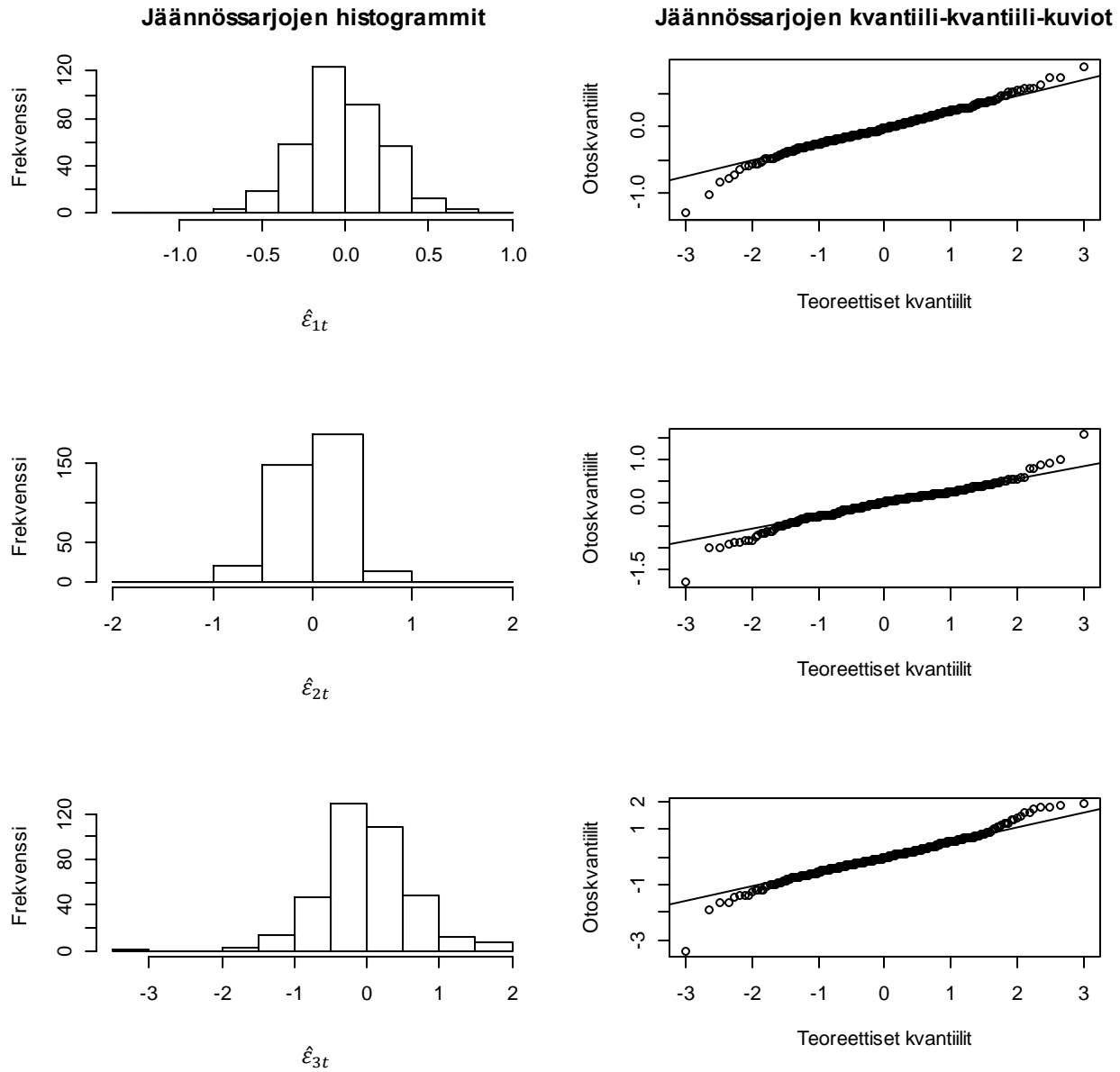
Kuvio 10. Otosautokorrelaatiofunktiot kolmelle estimoidulle faktorille $\hat{\beta}_{1t}$, $\hat{\beta}_{2t}$ ja $\hat{\beta}_{3t}$ sekä jäännössarjoille $\hat{\varepsilon}_{1t}$, $\hat{\varepsilon}_{2t}$ ja $\hat{\varepsilon}_{3t}$ käyttäen AR(2)-mallia.

Vaikka kuvion 10 (oikea sarake) estimoitujen faktoreiden $\hat{\beta}_{1t}$, $\hat{\beta}_{2t}$ ja $\hat{\beta}_{3t}$ jäännössarjat $\hat{\varepsilon}_{1t}$, $\hat{\varepsilon}_{2t}$ ja $\hat{\varepsilon}_{3t}$ näyttävät AR(2)-mallin sovittamisen jälkeen olevan stationaarisia, tehdään kuitenkin vielä vaadittavat testit, jotta nähdään, onko malli riittävä. Taulukossa 4 esitetävän Ljung-Box-testin mukaan jokaiselle faktorille sovitettu AR(2)-malli on riittävä, sillä hypoteesi jäännössarjan valkoisuudesta hyväksytään. Box-Ljung testi on laskettu kuvion 10 mukaisesti viivepituudelle 60. Samaan päätelmään päädytään myös käännepistetestissä. Laajennetun Dickey-Fuller-yksikköjuuritestin mukaan faktorilla $\hat{\beta}_{2t}$ saattaa olla yksikköjuuri, kun taas faktoreilla $\hat{\beta}_{1t}$ ja $\hat{\beta}_{3t}$ ei ole yksikköjuurta. Testin kriittiset arvot hylkäämi-

selle ovat -3,98 yhden prosentin tasolla, -3,42 viiden prosentin tasolla ja -3,13 kymmenen prosentin tasolla. Jarque-Bera-testi kuitenkin hylkää hypoteesin jäännössarjojen normaalisuudesta. Samaan päätelmään tullaan tarkastelemalla kvantiili-kvantiili-kuvioita (kuvio 11 oikea sarake), jonka perusteella aineiston havainnot ovat liian suuria verrattuna normaalijakaumaan. Kuviosta nähdään, että erityisesti jäännössarjan $\hat{\varepsilon}_{3t}$ jakauma on oikealle vino. Kaikkien kolmen jäännössarjan kohdalla myös pienimmät havainnot ovat liian pieniä verrattuna normaalijakaumaan, joten jakauma on paksuhäntäinen molempiin suuntiin. Tämä voidaan todeta myös histogrammeista kuviosta 11 (vasen sarake).

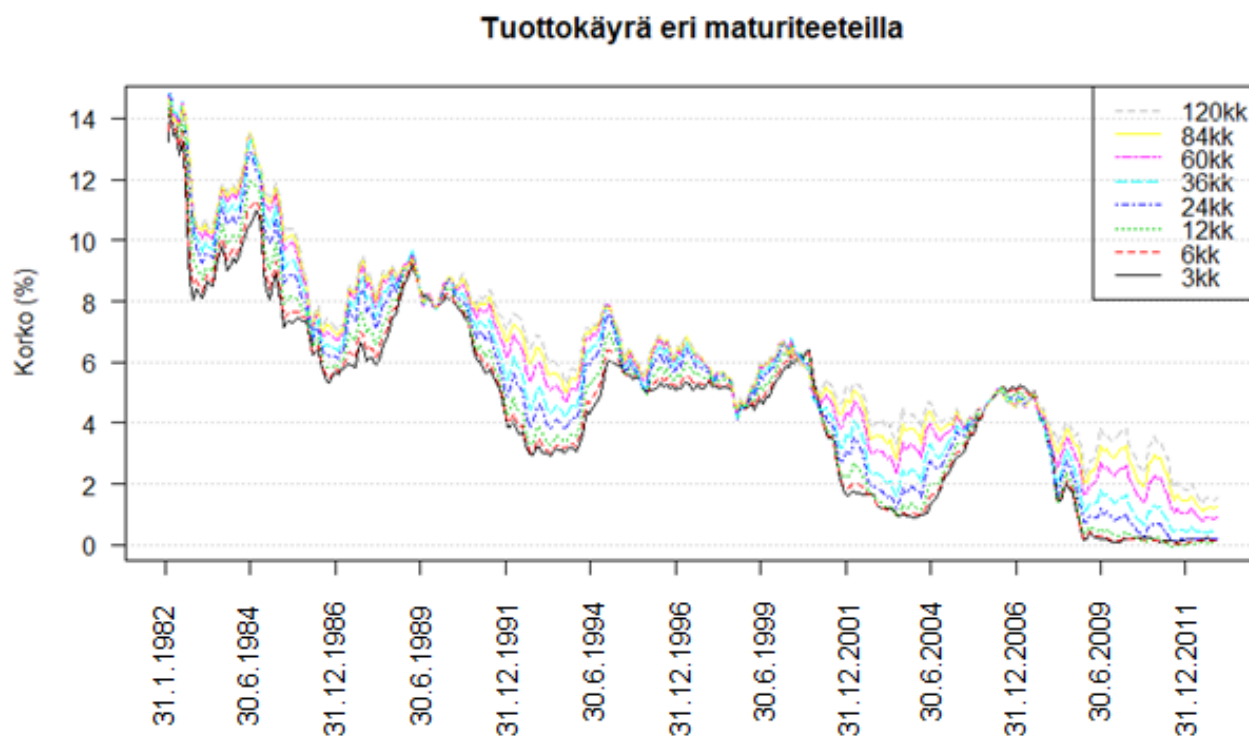
Taulukko 4. Sovitetun Nelson-Siegel -tuottokäyrän faktoreiden $\hat{\beta}_{1t}$, $\hat{\beta}_{2t}$ ja $\hat{\beta}_{3t}$ ja jäännössarjojen $\hat{\varepsilon}_{1t}$, $\hat{\varepsilon}_{2t}$ ja $\hat{\varepsilon}_{3t}$ valkaisuuden testaamiseen perustuvia arvoja, kun $\lambda_t = 0,0609$. Ensimmäinen sarake sisältää Box-Ljung-testin tulokset jäännössarjoille, toinen laajennetun Dickey-Fuller -yksikköjuuritestin (ADF) arvot faktoreille, kolmas käännepestetestin vastaavat arvot jäännössarjoille ja neljäs Jarque-Bera-testin tulokset jäännössarjoille.

Faktori	Box-Ljung	Käännepestesteti	ADF	Jarque-Bera
$\hat{\beta}_{1t}$	0,949	0,537	-4,407	4,921e-12
$\hat{\beta}_{2t}$	0,347	0,537	-3,324	< 2,2e-16
$\hat{\beta}_{3t}$	0,101	0,622	-4,573	< 2,2e-16



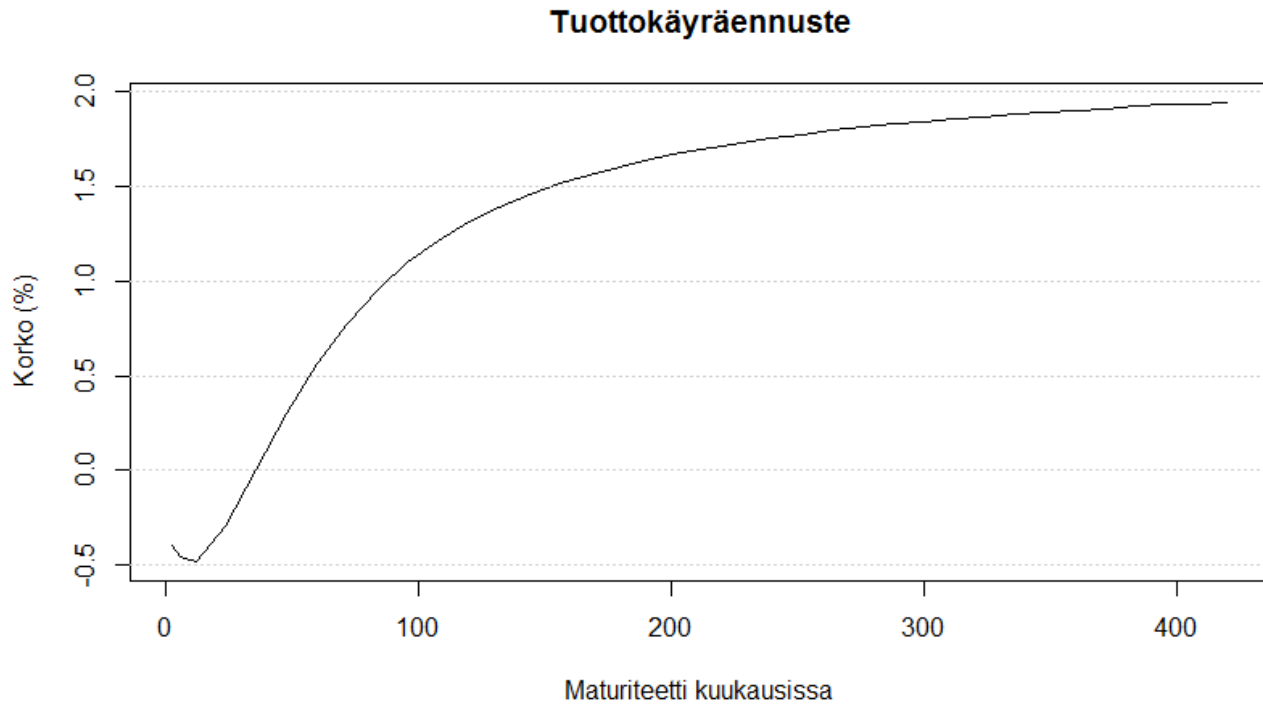
Kuvio 11. Estimoitujen faktoreiden $\hat{\beta}_{1t}$, $\hat{\beta}_{2t}$ ja $\hat{\beta}_{3t}$ jäännössarjojen $\hat{\varepsilon}_{1t}$, $\hat{\varepsilon}_{2t}$ ja $\hat{\varepsilon}_{3t}$ jakaumat.

Kuvio 12 näyttää sovitetun korkomallin markkinatuotot Yhdysvaltojen liikkeelle laskemista velkakirjoista eri maturiteeteilla. Ajan myötä laskeva trendi on selkeästi havaittavissa ja tuottoerot eri maturiteettien välillä ovat laajat, kun tuotot putoavat jyrkästi alaspäin. Kuvio käsittää 31 vuoden tuotot ja siitä nähdään selkeästi, että mitä pitempiaikainen maturiteetti on, sitä korkeampi on tuotto.



Kuvio 12. Sovitetun korkomallin markkinatuotot Yhdysvaltojen liikkeelle laskemista arvopapereista eri maturiteeteilla.

Dynaamista Nelson-Siegel-mallia voidaan käyttää tulevien tuottokäyrien ennustamiseen. Käytetyssä korkoaineistossa maturiteetteja on kahdeksan aina kolmesta kuukaudesta kymmeneen vuoteen asti. Jotta myöhemmin tutkielmassa suoritettavat eläkkeiden esimerkkilaskelmat voidaan toteuttaa, on tuottokäyrää ensin ekstrapoloitava tulevaisuuteen. Kuviossa 13 nähtävä tuottokäyräennuste perustuu mallinnetun tuottokäyrän simulointiin. Simuloinnissa on käytetty R:n sisäistä funktiota *arima.sim* ja simulointi on toteutettu vuoden päähän ennustetuista tuotoista. Tutkielmassa myöhemmin toteutettavissa eläkkeiden esimerkkilaskelmissa tullaan käyttämään kyseistä ekstrapoloitua tuottokäyrää.



Kuvio 13. Tuottokäyrään eli aikasarjaestimaatteihin $\hat{\beta}_{1t}$, $\hat{\beta}_{2t}$ ja $\hat{\beta}_{3t}$ perustuva ennuste vuoden päähän. Ennuste perustuu mallinnetun tuottokäyrän simulointiin eri maturiteeteilla.

Vuosikymmenten ajan korkojen positiivisuus on ollut yksi reunaehto järkeväksi katsotuille korkomalleille. Tämän vuoksi on muun muassa suosittu mieluummin log-normaaleja kuin normaalijakautuneita malleja korkotuotoille. Äskettäin kuitenkin Euroopan keskuspankki (EKP) teki jotain, mitä yksikään suuri keskuspankki ei ole aiemmin kokeillut ja painoi pankkien lyhyen talletuskoron negatiiviseksi EU:ssa. Täten EKP pyrkii saamaan vauhtia euroalueen hauraaseen talouteen sekä kiihdyttämään alhaiseksi painunutta inflaatiota. (Paavonen, 2014.) Kuvion 13 tuottokäyrät näyttäisivät ajautuvan negatiiviseksi lyhytaikaisimpien maturiteettien kohdalla etenkin mitä pidemmälle tuottoja pyritään ennustamaan. Korkoaineiston ja siihen sovitetun mallin ominaisuuksia (katso kuvio 12) tarkastelemalla voidaan todeta, että stokastinen malli voi ekstrapoloinnin sekä satunnaisvaihtelun vuoksi ajaa ennusteita negatiivisiin arvoihin. Päätelmää tukee myös edellä mainittu EKP:n viimeaikaiset korkopäätökset. Eläkevakuutuksen vakuutusmaksun sekä vastuvelan määrittely perustuu kassavirtojen nykyarvoihin. Teoreettisesti ajatellen negatiivisten arvojen säilyttäminen mallissa sisältää ongelmia, sillä kun lasketaan kassavirtojen nykyarvoa, ajatellaan yleisesti, että korkojen on oltava positiivisia. Tämä ajattelutapa pienentää vakuutusmaksua sekä vastuvelkaa. Näitä vastaavat varat vakuutusyhtiö sijoittaa muun muassa osake-

ja korkomarkkinoille ja saa niille vuosittain tuottoa, jonka tulisi olla vähintään diskonttauskoron suuruinen. Tutkielman esimerkkilaskelmissa negatiiviset korot asetetaan nolliksi, jotta ne eivät aiheuta suurempia ongelmia pääoma-arvojen muodostamisessa.

5.3 Elinikäisen eläkkeen pääoma-arvot

Tässä luvussa katsotaan, miten muodostetut kuolevuus- ja korkomalli ovat yhteydessä toisiinsa eri tilanteissa. Tähän mennessä on onnistuttu saamaan vain vähän todisteita kuolevuuden ja taloudellisten tekijöiden riippuvuudelle (Ronkainen, 2012). Seuraava haaste onkin kehittää sellainen joustava malli riippuvuuksien rakenteen mallintamiseen, joka mahdollistaa stokastisten skenaarioiden generoimisen. Aluksi pääoma-arvojen jakauma muodostetaan ilman kuolevuutta, minkä jälkeen kuolevuus huomioidaan laskelmissa ja jakauma muodostetaan samoin kuin ilman kuolevuutta.

Tutkielmassa muodostetaan elinikäisen eläkkeen pääoma-arvojen jakaumat vakuutustuotteelle ilman kuolemanvaravakuutusta. Kuolemanvaravakuutuksella tarkoitetaan sellaista korvaussummaa, jonka yhtiö suorittaa kuolinhetkellä, jos tämä tapahtuu vakuutuskauden aikana (Nyrhinen, 2012). Tarkoituksena on simuloida diskontatuille elinkoroille todennäköisyysjakauma sekä tutkia eläkerahaston vakavaraisuutta. Ajatuksena on valita tietyn ikäinen ja kokoinen joukko eläkeläisiä ja ennustaa heidän eläke-menonsa eli maksettavien eläkkeiden määrä stokastisen mallin avulla korkojen diskonttaus ja kuolevuus huomioiden.

Esimerkkilaskelmissa elinikäisen eläkkeen pääoma-arvoja lasketaan kohortille, jossa eläkeläiset ovat iältään täsmälleen $x = 65$ perusvuoden $t = 0$ alussa. Tutkielmassa kuolevuusmallin perusteella oletetaan, että elämä päättyy viimeistään iässä $\omega = 100$. Valitaan myös kohortin kooksi $N_0 = 10\,000$. Muodostetaan aluksi esimerkkilaskelmat yllä mainituilla edellytyksillä ilman kuolevuutta, eli oletetaan, että koko kohortille maksetaan yksikköeläkettä 35 vuotta. Tällöin stokastisuutta pääoma-arvon laskentaan tulee ainoastaan korkojen kautta. Lopuksi laskelmat tehdään myös kuolevuus huomioiden.

Yleensä tulevat kassavirrat diskontataan raportointihetkelle käyttäen riskitöntä korkokäyrää. Yleisesti tällainen korkokäyrä on Euro Swap -korkokäyrä, joka ohjeistetaan Solvenssi II:n mukaan muuntamaan nollakuponkikäyräksi. Kuitenkin osa vakuutusyhtiöistä on käyttänyt riskittömänä korkokäyränä joko muunneltua swap-käyrää tai valtioiden liikkeelle laskemien joukkovelkakirjojen korkotasoihin perustuvaa korkokäyrää. (Saraste, 2009.) Tutkielman korkoaineisto eroaa alkuperältään Solvenssi II:ssa käy-

tettävästä aineistosta. Lisäksi Euroopassa on valuutoita, joihin sovelletaan valtionlainoista johdettuja korkoja swap-markkina-aineiston riittämättömyyden vuoksi, joten laskelmissa on perusteltua käyttää simuloitua aineistoa sellaisenaan. Päätöstä tukee myös se, että Yhdysvaltojen valtion lainojen korko on teoriassa lähempänä riskitöntä korkoa kuin yritysten välisten swap-markkinoiden korko. Tämä on myös yksi syy, miksi korkoja korjataan jollakin kertoimella Solvenssi II:ssa.

Taulukossa 5 esitetään elinikäisen eläkkeen pääoma-arvoja. Pääoma-arvojen jakaumalla PV_a tarkoitetaan jakaumaa, jossa ei ole otettu huomioon kuolevuutta. Jakauman odotusarvoa $E(PV_a) = 25,23$ käytetään yleensä teoreettisena nettovakuutusmaksuna. Pääoma-arvojen jakaumassa PV_b kuolevuus taas on mukana laskelmissa. Teoreettiseksi nettovakuutusmaksuksi saadaan odotusarvo $E(PV_b) = 18,45$.

Taulukko 5. Elinikäisen eläkkeen pääoma-arvoja PV/N_0 . Pääoma-arvot on laskettu sekä ilman kuolevuutta PV_a että kuolevuus huomioiden PV_b .

Pääoma-arvo	Minimi	Keskiarvo	Maksimi
PV_a	7,23	25,23	34,00
PV_b	6,64	18,45	23,39

5.4 VaR-luku

VaR-lukuja lasketaan yleensä simuloidusta aineistosta, koska simulointien määrää kasvattamalla saadaan havaintoja myös kaukaa jakauman häntäalueelta, jolloin VaR-luvuista tulee luotettavampia. Valitaan Solvenssi II:n mukainen 99,5 prosentin kvantiili ja lasketaan taulukossa 5 esitetyille pääoma-arvojen jakaumille VaR-luvut käyttäen R-ohjelmiston funktiota *quantile* löytämään pääoma-arvojen vaihtelua vastaava kvantiili. Tällöin jakaumia vastaaviksi arvoiksi saadaan $\text{VaR}(PV_a) = 34,00$ ja $\text{VaR}(PV_b) = 23,39$. VaR-luvuilla tarkoitetaan sitä, että vakuutusyhtiön tulee varata vähintään kyseinen määrä, jotta tulevat eläkkeet pystytään 99,5 prosentin todennäköisyydellä kattamaan.

Muodostetaan seuraavaksi diskontatut jakaumat vakuutusyhtiön oman varallisuuden OF muutokselle vuoden aikana kaavan (4.9) mukaisesti ilman kuolevuutta ($OF(t_a)$) ja kuolevuus huomioiden ($OF(t_b)$). Valitaan mukaan maturiteetit 36, 60, 84 ja 120. Taulukossa 6 esitetään VaR-lukuja vakuutusyhtiön oman varallisuuden OF muutokselle vuoden aikana eripituisille korkosijoituksille. Taulukos-

ta 6 nähdään, että mitä pidempiaikainen sijoitus on, sitä pienempi on pääomavaatimus. Korko on usein riippuvainen laina-ajan pituudesta, jolloin pitkille talletuksille maksetaan korkeampaa korkoa. Tämä johtuu siitä, että pidempiaikaiset korot vaihtelevat vähemmän sekä ovat tuotoiltaan korkeampia kuin lyhyen maturiteetin korkosijoitukset. Lisäksi pääomavaatimus näyttäisi olevan selkeästi alhaisempi, kun kuolevuus otetaan mukaan. Tämä johtuu eläkettä nauttivien eläkeläisten määrän vähenemisestä ajan mittaan.

Taulukko 6. VaR-lukuja eripituisille korkosijoituksille. Luvut on laskettu sekä ilman kuolevuutta $OF(t_a)$ että kuolevuus huomioiden $OF(t_b)$.

Maturiteetti	$OF(t_a)$	$OF(t_b)$
36	94453,30	69280,41
60	82724,48	47957,10
84	71504,64	27813,20
120	52201,08	6281,35

6 JOHTOPÄÄTÖKSET

On selvää, että naiset ja miehet on jaettava omiksi populaatioikseen kuolevuutta tutkittaessa. Tutkielmassa päädyttiin tutkimaan kuolevuutta vain naisten osalta, sillä varsinkin miesten kohdalla eläkevakutuskuolevuuteen vaikuttaa varallisuus. Naisten kohdalla samankaltaista riippuvuutta ei ole. Tutkielmassa muodostettu Makehamin malli onnistuu kuvaamaan kuolevuutta melko tarkasti naisten ryhmäläkkeen osalta.

Korkoriskit ovat henkivakuutusyhtiöille usein hyvin merkittäviä ja niitä voidaan mitata monella eri tapaa, kuten esimerkiksi Solvenssi II:n mukaisissa vakavaraisuuskehikoissa. Tutkielmassa muodostettiin dynaaminen Nelson-Siegel-korkomalli kuvaamaan Yhdysvaltojen liikkeelle laskemien arvopapereiden markkinatuottoja kahdeksalla eri maturiteetilla. Korkomalli onnistuu toistamaan havaitun tuottokäyrän eri muotoja hyvin, mutta sovitetun mallin käyrillä on kuitenkin ongelmia sellaisten päivämäärien kohdalla, jolloin tuotot ovat hajanaisia eri maturiteeteilla. Kaiken kaikkiaan malli kuitenkin näyttäisi sopivan havaittuun aineistoon hyvin.

Eläkevakuutuksen keskeisimmät riskit johtuvat pitkäikäisyyden sekä sijoitusten tuottojen kehityksestä. Ei kuitenkaan riitä tarkastella vain pitkäikäisyyttä, kun ollaan tekemisissä eläkevakuutuksen riskienhallinnan kanssa vaan mukaan tarkasteluun on otettava myös pitkäikäisyysriskin rahoittaminen. Tutkielmassa lasketut elinikäisen eläkkeen pääoma-arvot valitulle kohortille liittyvät yhteen aiemmin muodostetun kuolevuus- sekä korkomallin. Laskemalla odotusarvo pääoma-arvojen jakaumalle ilman kuolevuutta, saadaan teoreettiseksi nettovakuutusmaksuksi 25,23 ja kuolevuuden kanssa 18,45.

Lisäksi Solvenssi II:n mukaisen SCR-pääomavaatimuksen tarkastelussa laskettiin joitakin VaR-lukuja vakuutusyhtiön oman varallisuuden $OF(t)$ muutokselle vuoden aikana eripituisille korkosijoituksille. Laskelmat toteutettiin erikseen aineistolle, jossa kuolevuutta ei oteta huomioon sekä kuolevuuskehityksen huomioivalle aineistolle. Korkoaineiston ja siihen sovitetun mallin ominaisuuksista johtuen stokastinen malli ajaa ennusteita negatiivisiin arvoihin, jolloin osa simuloidun ennusteen koroista oli negatiivisia. Esimerkkilaskelmissa negatiiviset korot asetettiin nolliksi. VaR-lukuja laskemalla havaitaan, että mitä pidempiaikainen sijoitus on, sitä pienempi on pääomavaatimus. Kuolevuudella näyttäisi lisäksi olevan alentava vaikutus pääomavaatimukseen.

LÄHTEET

- Aho, J. & Sankala, M. (2011), *Pääoma-arvokertoimet*, Eläketurvakeskus. Saatavilla Internetistä: http://www.etk.fi/fi/gateway/PTARGS_0_2712_459_440_3034_43/http%3B/content.etk.fi%3B7087/publishedcontent/publish/etkfi/fi/julkaisut/k%C3%A4sikirjat/paaomaarvokertoimet_7.pdf. Viitattu 4.5.2014.
- Brockwell, P.J. & Davis, R.A. (2002), *Introduction to Time Series and Forecasting, Second Edition*, Springer-Verlag, 45-110, 138-178.
- Cryer, J.D. & Chan, K-S. (2008), *Time Series Analysis with Applications in R, Second Edition*, Springer, 55-86.
- Diebold, F.X. & Li, C. (2006), *Forecasting the term structure of government bond yields*, Journal of Econometrics, 130, 337-364.
- Diebold, F.X. & Rudebusch, G.D. (2013), *Yield Curve Modeling and Forecasting – The Dynamic Nelson-Siegel Approach*, Princeton University Press, 4-54.
- Doray, L.G. (2008), *Inference for Logistic-type Models for the Force of Mortality*, The Society of Actuaries, 3-6.
- Eläketurvakeskus (2014), *Työnantajien järjestämä kollektiivinen lisäeläketurva*. Saatavilla Internetistä: http://www.etk.fi/fi/service/tyonantajien_jarjestama/266/tyonantajien_jarjestama. Viitattu 13.2.2014.
- Finanssivalvonta (2012), *Sukupuolen käyttäminen vakuutusmaksujen tai vakuutuksesta saatavien etuuksien laskennassa*, Saatavilla Internetistä: http://www.finanssivalvonta.fi/fi/Finanssiasiakas/Tuotteita/Vakuutukset/Sukupuoli_laskennassa/Pages/Default.aspx#.U9EnkrGWau5. Viitattu 24.7.2014.
- Finanssivalvonta (2014), *Solvenssi II -hanke*. Saatavilla Internetistä: <http://www.finanssivalvonta.fi/fi/Saantely/Saantelyhankkeet/Solvenssi/Pages/Default.aspx>. Viitattu 13.5.2014.
- Forfar, D. (2004), *Mortality Laws*. Saatavilla Internetistä: http://maths.swan.ac.uk/staff/zs/mortality_laws.pdf. Viitattu 15.3.2014.
- Fox, J. (2002), *Nonlinear Regression and Nonlinear Least Squares*. Saatavilla Internetistä: <http://cran.r-project.org/doc/contrib/Fox-Companion/appendix-nonlinear-regression.pdf>. Viitattu 20.3.2014.
- Graybill, F.A. & Iyer, H.K. (1994), *Regression Analysis: Concepts and Application*, Duxbury Press, 86-87.

- Hastie, T., James, G., Tibshirani, R. & Witten, D. (2013), *An Introduction to Statistical Learning with Applications in R*, Springer. 82-92.
- Jorion, P. (1997), *Value at Risk: The New Benchmark for Controlling Market Risk*, McGraw-Hill, 99.
- Kauppi, M. (2004), *Lakisääteisen tapaturmavakuutuksen referenssikuolevuus*, Tapaturmavakuutuslaitosten liitto.
- Kolehmainen, P. (2009), *Vapaamuotoisen työeläkevakuutuksen kuolevuus*, SHV-harjoitustyö.
- Korhonen, S. (2009), *Työeläkeyhtiöiden joukkovelkakirjasalkun hallinta*, Pro gradu -tutkielma. Kansantaloustiede. Tampere: Tampereen yliopisto.
- Malz, A.M. (2011), *Financial Risk Management: Models, History, and Institutions*, Wiley Finance, 93-113.
- Mandatum Life (2012), *Eläketurvan täydentäminen*. Saatavilla Internetistä: <https://www.mandatumlife.fi/palvelut/tuotteet/elaketurvan-taydentaminen>. Viitattu 27.6.2014.
- Martelin, T., Koskinen, S. & Valkonen, T. (2005), *Suomalaisten terveys - Kuolleisuus*, Kustannus Oy Duodecim. Saatavilla Internetistä: http://www.terveyskirjasto.fi/terveyskirjasto/tk.koti?p_artikkeli=suo00021. Viitattu 22.7.2014.
- Mellin, I. (2010), *Aikasarja-analyysi*, Aalto-yliopisto.
- Mäkinen, M. (2004), *Referenssikuolevuus henkivakuutusyhtiöille*, Suomen Aktuaariyhdistys.
- Nyrhinen, H. (2012), *Henkivakuutusmatematiikka*, Helsingin yliopisto.
- Paavonen, A. (2014), *EKP teki historiaa – laski molempia korkoja*, Taloussanomat. Saatavilla Internetistä: <http://www.taloussanomat.fi/rahoitus/2014/06/05/ekp-teki-historiaa-laski-molempia-korkoja/20147973/12>. Viitattu 25.6.2014.
- Pesonen, M., Soininen, P. & Tuominen, T. (1999), *Henkivakuutusmatematiikka*, Suomen vakuutusalan koulutus ja kustannus Oy, 38–47.
- Ronkainen, V. (2012), *Stochastic modeling of financing longevity risk in pension insurance*, Finanssi-valvonta. Väitöskirja. Saatavilla Internetistä: <http://www.suomenpankki.fi/fi/julkaisut/tutkimukset/erillisjulkaisut/Pages/default.aspx>. Viitattu 9.5.2014.
- R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Saatavilla Internetistä: <http://www.R-project.org/>.
- Saraste, L. (2009), *Henkivakuutusyhtiön korkoriski, sen suojaus ja vaikutukset vakavaraisuuteen*, Suomen aktuaariyhdistys.
- Seber, G.A.F. & Wild, C.J. (1989), *Nonlinear Regression*, John Wiley & Sons, 4-10, 21-23.

SHY (1982), *Henki- ja eläkevakuutuksen vakuutustekniikka*.

Sirén, T. (2012), *Henkivakuutusyhtiöiden kuolevuustutkimus K2012*, Finanssivalvonta. Saatavilla Internetistä: http://www.finanssivalvonta.fi/fi/Tiedotteet/Verkkouutiset/Documents/Henkivakuutusyhtioiden_kuolevuustutkimus_K2012.pdf. Viitattu 13.2.2014.

Suomen virallinen tilasto (SVT), *Kuolleet*. Saatavilla Internetistä: <http://tilastokeskus.fi/til/kuol/index.html>. Viitattu 7.3.2014

Tiainen, S. (2011), *Suomalaisten vakuutusyhtiöiden Solvenssi II -vaikuttavuustutkimus*, Finanssivalvonta. Saatavilla Internetistä: <http://www.finanssivalvonta.fi/fi/Tiedotteet/Lehdistotiedotteet/Documents/QIS5.pdf>. Viitattu 17.5.2014

Tsay, R.S. (2010), *Analysis of Financial Time Series, Third Edition*, John Wiley & Sons, 325-338.

Ufer, W. (1996), *The “Value at Risk” Concept for Insurance Companies*, *actuaries.org*. Saatavilla Internetistä: <http://www.actuaries.org/AFIR/Colloquia/Nuernberg/Ufer.pdf>. Viitattu 25.7.2014.

Vedenjuoksu, T. (2011), *Talousmatematiikka*, Oulun yliopisto.

LIITTEET

Liite A: Makeham-funktiosovite

Funktiosovitteen tekemiseen käytetty Makeham-funktio. Aineistona toimii $2 \cdot 10^1$ kokoinen 'data' -niminen taulu. Taulun kaksi saraketta kuvaavat ikää (x) sekä kyseisen iän kuolevuutta (y).

```
x <- c(0:100)
## Logaritmoidaan y:
y <- log(y)

min.fun <- function(param,x,y){
  a <- param[1]
  b <- param[2]
  c <- param[3]
  if(c<0) return(100000)
  sum((y-log(a+b*exp(c*x)))^2)}

## par0:aan sopivat alkuarvot i, j:
lm(y~x)
par0 <- c(0,exp(i),j)
## Otetaan mukaan vain yli 40-vuotiaat:
result.k <- nlm(min.fun,par0,x[41:100],y[41:100], hessian=TRUE, typsize=par0)
## Funktiosovitteen kertoimet a, b ja c:
result.k$estimate

## Kuolevuusfunktion kuvaaja:
plot(x,y)
a <- result.k$estimate[1]
b <- result.k$estimate[2]
c <- result.k$estimate[3]
yhat <- log(a+b*exp(c*x))
lines(x,yhat)
```

Liite B: DNS-korkomalli

Korkomallin tekemiseen käytetty dynaaminen Nelson-Siegel-malli. Aineistona toimii $8 \cdot 371$ kokoinen 'data2' -niminen taulu. Taulun kahdeksan saraketta kuvaavat eri maturiteettien tuottoa annetulla päivämäärällä.

```
y <- c(0:370)
x <- matrix(y,nrow=8,ncol=371)
## Maturiteetit kuukausissa:
mat <- c(3,6,12,24,36,60,84,120)
## Vakio lambda sekä mallin kertoimet:
lambda<-0.0609
a <- (1-exp(-lambda*mat))/(lambda*mat)
b <- ((1-exp(-lambda*mat))/(lambda*mat))-exp(-lambda*mat)

for (i in 1:length(x)) {
  tulos <- lm(x~a+b)
}
## Antaa parametrit beta1, beta2 ja beta3, kun lambda=0.0609:
parametrit <- t(tulos$coef)
```